VISIT…



The Best Damn Cybercrime and Digital Forensics Book Period

Kevin Cardwell
Timothy Clinton
Tyler Cohen
Edward Collins
James "Jim" Cornell
Michael Cross
Larry Depew
Art Ehuan
Michael Gregg
Captain Benjamin R. Jean
Kevin O'Shea
Kevin Reis
Anthony Reyes
Sondra Schneider
Amber Schroader
Karen Schuler
Jesse Varsalone
Jack Wiles
Craig Wright

Elsevier, Inc., the author(s), and any person or firm involved in the writing, editing, or production (collectively "Makers") of this book ("the Work") do not guarantee or warrant the results to be obtained from the Work.

There is no guarantee of any kind, expressed or implied, regarding the Work or its contents. The Work is sold AS IS and WITHOUT WARRANTY. You may have other legal rights, which vary from state to state.

In no event will Makers be liable to you for damages, including any loss of profits, lost savings, or other incidental or consequential damages arising out from the Work or its contents. Because some states do not allow the exclusion or limitation of liability for consequential or incidental damages, the above limitation may not apply to you.

You should always use reasonable care, including backup and other appropriate precautions, when working with computers, networks, data, and files.

Syngress Media®, Syngress®, "Career Advancement Through Skill Enhancement®," "Ask the Author UPDATE®," and "Hack Proofing®," are registered trademarks of Elsevier, Inc. "Syngress: The Definition of a Serious Security Library"™, "Mission Critical™," and "The Only Way to Stop a Hacker is to Think Like One™" are trademarks of Elsevier, Inc. Brands and product names mentioned in this book are trademarks or service marks of their respective companies.

KEY	SERIAL NUMBER
001	HJIRTCV764
002	PO9873D5FG
003	829KM8NJH2
004	VBPK567AQT
005	CVPLQ6WQ23
006	VBP965T5T5
007	HJJJ863WD3E
008	2987GVTWMK
009	629MP5SDJT
010	IMWQ295T6T

PUBLISHED BY
Syngress Publishing, Inc.
Elsevier, Inc.
30 Corporate Drive
Burlington, MA 01803

The Best Damm Cybercrime and Digital Forensics Book Period

Copyright © 2007 by Elsevier, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher, with the exception that the program listings may be entered, stored, and executed in a computer system, but they may not be reproduced for publication.

Printed in the United States of America
1 2 3 4 5 6 7 8 9 0

ISBN 13: 978-1-59749-228-7

Publisher: Amorette Pedersen
Acquisitions Editor: Andrew Williams
Cover Designer: Michael Kavish
Page Layout and Art: SPi

For information on rights, translations, and bulk sales, contact Matt Pedersen, Commercial Sales Director and Rights, at Syngress Publishing; email m.pedersen@elsevier.com.



Contributing Authors

Kevin Cardwell (CEH, ECSA, LPT) works as a freelance consultant and provides consulting services for companies throughout the U.S., U.K., and Europe. He is an adjunct associate professor for the University of Maryland University College, where he participated in the team that developed the Information Assurance Program for Graduate Students, which is recognized as a Center of Excellence program by the National Security Agency (NSA). He is an instructor and technical editor for computer forensics and hacking courses. He has presented at the Blackhat USA Conference.

During a 22-period in the U.S. Navy, Kevin tested and evaluated surveillance and weapon system software. Some of this work was on projects like the Multi-Sensor Torpedo Alertment Processor (MSTRAP), Tactical Decision Support System (TDSS), Computer Aided Dead Reckoning Tracer (CADRT), Advanced Radar Periscope Discrimination and Detection (ARPDD), and the Remote Mine Hunting System (RMHS). He has worked as both a software and systems engineer on a variety of Department of Defense projects and was selected to head the team that built a Network Operations Center (NOC) that provided services to the command ashore and ships at sea in the Norwegian Sea and Atlantic Ocean. He served as the leading chief of information security at the NOC for six years prior to retiring from the U.S. Navy. During this time he was the leader of a five-person Red Team.

Kevin wishes to thank his mother, Sally; girlfriend, Loredana; and daughter, Aspen, all of whom are sources of his inspiration. Kevin holds a master's degree from Southern Methodist University and is a member of the IEEE and ACM. Kevin currently resides in Cornwall, England.

Timothy Clinton has held multiple roles in the EDD/ESI vendor space. He is currently employed as forensics operations manager for the National Technology Center division of Document Technologies, Inc. (DTI), a major ESI service. Since joining the DTI team, Mr. Clinton has served in multiple roles, including EDD production manager, technical architect,

and forensic investigator. He has conducted and managed investigations for numerous civil cases regarding matters for Fortune 50 of law. Mr. Clinton's most notable achievement while at DTI is being responsible for the design and implementation of a showcase data forensics laboratory in Atlanta, Georgia.

Tyler Cohen (CISSP) is employed by Computer Science Corporation contracted as a researcher and developer for the Department of Defense Cyber Crime Center. Her specialty is digital forensics and intrusions. She is considered an expert in hacking and conducting forensic exams with the iPod and other alternative media devices. She presents her expertise at various conferences all over the country some of which include the Department of Defense Cyber Crime Conference, International High Technology Crime Investigation Association and The California District Attorney's Cyber Crime Conference.

Edward Collins (CISSP, CEH, Security+, MCSE:Security, MCT) is a senior security analyst for CIAN, Inc., where he is responsible for conducting penetration tests, threat analysis, and security audits. CIAN (www.ciancenter.com) provides commercial businesses and government agencies with all aspects of information security management, including access control, penetration testing, audit procedures, incident response handling, intrusion detection, and risk management. Edward is also a training consultant, specializing in MCSE and Security+ certifications. Edward's background includes positions as information technology manager at Aurora Flight Sciences and senior information technology consultant at Titan Corporation.

James "Jim" Cornell (CFCE, CISSP, CEECS) is an employee of Computer Sciences Corp. (CSC) and an instructor/course developer at the Defense Cyber Investigations Training Academy (DCITA), which is part of the Defense Cyber Crime Center (DC3) in Maryland. At the academy he teaches network intrusions and investigations, online undercover techniques, and advanced log analysis. He has over 26 years of law enforcement and over 35 years of electronics and computer experience. He is a member/coach of the International Association of Computer Investigative Specialists (IACIS) and a member of the International Information Systems Forensics Association (IISFA) and the International

Information Systems Security Certification Consortium (ISC2). He is currently completing the Certified Technical Trainer (CTT+) process and is a repeat speaker at the annual Department of Defense Cyber Crime Conference.

He would like to thank his mother for more than he can say, his wife for her patience and support, and Gilberto for being the best friend ever.

Michael Cross (MCSE, MCP+I, CNA, Network+) is an internet specialist/programmer with the Niagara Regional Police Service. In addition to designing and maintaining the Niagara Regional Police's Web site (www.nrps.com) and intranet, he has also provided support and worked in the areas of programming, hardware, database administration, graphic design, and network administration. In 2007, he was awarded a Police Commendation for work he did in developing a system to track high-risk offenders and sexual offenders in the Niagara Region. As part of an information technology team that provides support to a user base of over 1,000 civilian and uniformed users, his theory is that when the users carry guns, you tend to be more motivated in solving their problems.

Michael was the first computer forensic analyst in the Niagara Regional Police Service's history, and for five years he performed computer forensic examinations on computers involved in criminal investigations. The computers he examined for evidence were involved in a wide range of crimes, inclusive to homicides, fraud, and possession of child pornography. In addition to this, he successfully tracked numerous individuals electronically, as in cases involving threatening e-mail. He has consulted and assisted in numerous cases dealing with computer-related/Internet crimes and served as an expert witness on computers for criminal trials.

Michael has previously taught as an instructor for IT training courses on the Internet, Web development, programming, networking, and hardware repair. He is also seasoned in providing and assisting in presentations on Internet safety and other topics related to computers and the Internet. Despite this experience as a speaker, he still finds his wife won't listen to him.

Michael also owns KnightWare, which provides computer-related services like Web page design, and Bookworms, which provides online sales of merchandise. He has been a freelance writer for over a decade

and has been published over three dozen times in numerous books and anthologies. When he isn't writing or otherwise attached to a computer, he spends as much time as possible with the joys of his life: his lovely wife, Jennifer; darling daughter Sara; adorable daughter Emily; and charming son Jason.

Larry Depew, PMP, is the director of the New Jersey Regional Computer Forensic Laboratory (NJRCFL), a partnership between the FBI and State of New Jersey that provides forensic examinations and training to law enforcement in the field of digital forensics. He retired from the Federal Bureau of Investigation (FBI) as a supervisory special agent after nearly 32 years and is currently employed by the State of New Jersey. Larry leads a laboratory of 24 forensic examiners from nine law enforcement agencies servicing more than 550 federal, state, and local law enforcement agencies in New Jersey and the surrounding region.

Larry oversaw the overall construction of the NJRCFL's physical laboratory space and implemented a quality system for laboratory operations to meet client quality requirements for digital forensic examinations, law enforcement training, and expert testimony.

Prior to becoming director of the NJRCFL, Larry worked on several information technology projects at the FBI in Washington, D.C., including developing user requirements for case management systems, and as project manager for the deployment of the Investigative Data Warehouse (IDWv1.0). Larry is an experienced digital forensic examiner who has conducted more than 100 examinations and reviewed the output of more than 1,000 examinations performed by NJRCFL examiners. His digital forensic certifications include the FBI CART Forensic Examiner (Windows, Linux, and personal data assistants) and steganography investigator.

Larry chaired the FBI's Computer Analysis Response Team's (CART) first Standard Operating Procedure and Quality System committee, which formed the basis for today's RCFL National Program and CART Standard Operating Procedures.

Larry is an adjunct professor in digital forensics at The College of New Jersey (TCNJ). He has also taught digital forensics at the New Jersey Institute of Technology (NJIT). Larry is a project management professional certified through the Project Management Institute. He has

lectured at many government and private sector conferences on topics relating to data management, workflow, computer security, and digital forensics. He has appeared on the Fox network and the Philadelphia ABC affiliate as an expert regarding digital evidence and Internet safety. He has been interviewed by several national publications and regional newspapers regarding digital evidence analysis, computer security, and Internet safety.

Art Ehuan (CISSP, CFCE, EnCE) is a digital forensic expert with senior management experience in developing and implementing digital forensic facilities for corporations and the United States government.

Art previously managed the Information Security Department for USAA, a Fortune 200 financial services company, where he developed and implemented policies, process, and technology for a state-of-the-art digital forensic facility for handling computer forensics and electronic discovery. Art was previously the deputy chief information security officer at Northrop Grumman, where he developed and implemented three digital forensic facilities for the company. He also developed and implemented Cisco Systems' first digital investigative facility.

Art also has extensive government experience in digital forensics. He was formerly an FBI special agent certified as a Computer Analysis Response Team member and Air Force Office of Special Investigations special agent certified as a computer crime investigator.

Art was formerly an adjunct professor at Georgetown University, Duke University, and George Washington University, where he taught undergraduate and graduate courses on computer forensics, incident response, and computer crime.

Michael Gregg is the president of Superior Solutions, Inc. and has more than 20 years' experience in the IT field. He holds two associate's degrees, a bachelor's degree, and a master's degree and is certified as CISSP, MCSE, MCT, CTT+, A+, N+, Security+, CNA, CCNA, CIW Security Analyst, CCE, CEH, CHFI, CEI, DCNP, ES Dragon IDS, ES Advanced Dragon IDS, and TICSA.

Michael's primary duties are to serve as project lead for security assessments helping businesses and state agencies secure their IT resources and assets. Michael has authored four books, including: *Inside Network Security Assessment*, *CISSP Prep Questions*, *CISSP Exam Cram2*,

and *Certified Ethical Hacker Exam Prep2*. He also was the lead author for *Hack the Stack: Using Snort and Ethereal to Master the Eight Layers of an Insecure Network* (Syngress, ISBN: 9781597491099). He has developed four high-level security classes, including Global Knowledge's Advanced Security Boot Camp, Intense School's Professional Hacking Lab Guide, ASPE's Network Security Essentials, and Assessing Network Vulnerabilities. He has created over 50 articles featured in magazines and Web sites, including *Certification Magazine*, GoCertify, *The El Paso Times*, and SearchSecurity.

Michael is also a faculty member of Villanova University and creator of Villanova's college-level security classes, including Essentials of IS Security, Mastering IS Security, and Advanced Security Management. He also serves as a site expert for four TechTarget sites, including SearchNetworking, SearchSecurity, SearchMobileNetworking, and SearchSmallBiz. He is a member of the TechTarget Editorial Board.

Captain Benjamin R. Jean has spent his entire law enforcement career in the State of New Hampshire, starting in 1992 for the Deerfield Police Department. He is currently employed as a Law Enforcement Training Specialist for the New Hampshire Police Standards & Training Council and is Chief of the Training Bureau. Captain Jean teaches classes in various law enforcement topics, including computer crime investigation, and is an active member of the New Hampshire Attorney General's Cyber Crime Initiative. He was recently awarded the 2006 Cyber Crime Innovation Award and holds an Associate's Degree in Criminal Justice from New Hampshire Community Technical College and a Bachelor's Degree in Information Technology from Granite State College.

Kevin O'Shea is currently employed as a Homeland Security and Intelligence Specialist in the Justiceworks program at the University of New Hampshire. In this capacity, Mr. O'Shea supports the implementation of tools, technology, and training to assist law enforcement in the investigation of crimes with a cyber component. In one of Kevin's recent projects, he was a technical consultant and developer of a training program for a remote computer-forensics-viewing technology, which is now in use by the state of New Hampshire. He also has developed a computer-crime-investigative curriculum for the New Hampshire Police Standards and Training.

Kevin Reis (CISSP, CFE, GCFA, EnCE) has extensive public and private sector experience in the fields of computer forensics, network investigations, financial fraud investigations, and electronic discovery. Kevin began his career conducting counterintelligence investigations as a special agent with the Federal Bureau of Investigation (FBI), but he soon joined the FBI Computer Analysis Response Team (CART). As a CART field examiner, Kevin provided computer forensics support and technical consultation to investigations ranging from financial institution fraud and child pornography to espionage. Kevin then joined the National Aeronautics and Space Administration (NASA) Office of Inspector General (OIG) as a computer crime investigator (CCI), where he investigated computer and network intrusions at the Goddard Space Flight Center. Following his tenure at NASA, Kevin entered the private sector, working as a computer intrusion analyst at Aegis Research Corporation and then as a senior associate with the Forensic Technology Services practice of the Big Four accounting firm KPMG. While at KPMG, Kevin provided computer forensics, data analysis, e-discovery, and investigative services on financial fraud and civil litigation engagements. Following the events of September 11, 2001, Kevin reentered public service with the Department of Justice OIG as a special agent to build the OIG's computer forensics program. Kevin is currently a special agent with the Federal Deposit Insurance Corporation OIG Electronic Crimes Unit and a reserve Air Force Office of Special Investigations CCI.

Anthony Reyes is a retired New York City Police Department Computer Crimes Detective. While employed for the NYPD, he investigated computer intrusions, fraud, identity theft, child exploitation, intellectual property theft, and software piracy.

He was an alternate member of New York Governor George E. Pataki's Cyber-Security Task Force, and he currently serves as President for the High Technology Crime Investigation Association. He is the Education & Training Working Group Chair for the National Institute of Justice's Electronic Crime Partner Initiative. Anthony is also an Associate Editor for the *Journal of Digital Forensic Practice* and an editor for *The International Journal of Forensic Computer Science*.

He is an Adjutant Professor and is the Chief Executive Officer for the Arc Enterprises of New York, Inc. on Wall Street. Anthony has over 20 years

of experience in the IT field. He teaches for several government agencies and large corporations in the area of computer crime investigations, electronic discovery, and computer forensics. He also lectures around the world.

Karen Schuler is vice president of Consulting Practice Group. She brings over 15 years of management, technology, forensics, and electronic discovery experience to ONSITE^3's team of experts and specialists. Karen's experience ranges from the migration of data, enterprisewide technology planning and implementation, forensic investigations to large and complex litigation matters involving electronic discovery. As a former owner of a boutique computer forensics and security firm as well as a contracted computer forensic examiner for the U.S. Securities and Exchange Commission, she is an expert at understanding the intricate details involved in providing admissible and defensible evidence.

Karen has a wide range of experience in dealing with change management, technology assessments, and investigations as they relate to large corporate entities in the financial services industry, pharmaceutical, retail, manufacturing, health care, and technology fields. In addition, she has routinely been engaged on large, unwieldy electronic discovery projects where an expert is required to oversee the methodologies as well as provide recommendations for better practices.

Sondra Schneider is CEO and Founder of Security University, a Vienna, VA-based Qualified Computer Security and Information Assurance Training Company. For the past 18 years Sondra has been traveling around the world training network professionals to be network and security professionals. In 2004 she was awarded Entrepreneur of the Year at the First Annual Woman of Innovation Awards from the Connecticut Technology Council. She sits on the advisory board for three computer security technology companies and is a frequent speaker at computer security and wireless industry events. She is a founding member of the NYC HTCIA and IETF, and she works closely with ISC2, ISSA, and ISACA chapters and the vendor community to provide qualified computer security training and feedback. Sondra holds the CISSP, CEH, ECSA, LPT, and CHFI credentials.

Amber Schroader has been involved in the field of computer forensics for the past 17 years. Amber has developed and taught numerous training courses for the computer forensic arena, specializing in the field of

wireless forensics as well as mobile technologies. Amber is the CEO of Paraben Corporation and continues to act as the driving force behind some of the most innovative forensic technologies. As a pioneer in the field, Amber has been key in developing new technology to help investigators with the extraction of digital evidence from hard drives, e-mail, and handheld and mobile devices. Amber has extensive experience in dealing with a wide array of forensic investigators ranging from federal, state, local, and foreign government as well as corporate investigators. With an aggressive development schedule, Amber continues to bring new and exciting technology to the computer forensic community worldwide and is dedicated to supporting the investigator through new technologies and training services that are being provided through Paraben Corporation. Amber is involved in many different computer investigation organizations, including The Institute of Computer Forensic Professionals (ICFP) as the chairman of the board, HTCIA, CFTT, and FLETC.

Amber currently resides in Utah and Virginia with her two children, Azure and McCoy.

Jesse Varsalone (A+, Linux+, Net+, iNet+, Security+, Server+, CTT+, CIW Professional, CWNA, CWSP, MCT, MCSA, MSCE 2000/2003, MCSA/MCSE Security, MCSD, MCDBA, MCSD, CNA, CCNA, MCDST, Oracle 8i/9i DBA, Certified Ethical Hacker) is a computer forensic senior professional at CSC. For four years, he served as the director of the MCSE and Network Security Program at the Computer Career Institute at Johns Hopkins University. For the 2006 academic year, he served as an assistant professor of computer information systems at Villa Julie College in Baltimore, Maryland. He taught courses in networking, Active Directory, Exchange, Cisco, and forensics.

Jesse holds a bachelor's degree from George Mason University and a master's degree from the University of South Florida. He runs several Web sites, including mcsecoach.com, which is dedicated to helping people obtain their MCSE certification. He currently lives in Columbia, Maryland, with his wife, Kim, and son, Mason.

Jack Wiles is a security professional with over 30 years' experience in security-related fields, including computer security, disaster recovery, and physical security. He is a professional speaker and has trained federal agents, corporate attorneys, and internal auditors on a number of computer

crime-related topics. He is a pioneer in presenting on a number of subjects that are now being labeled "Homeland Security" topics. Well over 10,000 people have attended one or more of his presentations since 1988. Jack is also a cofounder and president of TheTrainingCo. He is in frequent contact with members of many state and local law enforcement agencies as well as special agents with the U.S. Secret Service, FBI, U.S. Customs, Department of Justice, the Department of Defense, and numerous members of high-tech crime units. He was also appointed as the first president of the North Carolina InfraGard chapter, which is now one of the largest chapters in the country. He is also a founding member and "official" MC of the U.S. Secret Service South Carolina Electronic Crimes Task Force.

Jack is also a Vietnam veteran who served with the 101st Airborne Division in Vietnam in 1967-68. He recently retired from the U.S. Army Reserves as a lieutenant colonel and was assigned directly to the Pentagon for the final seven years of his career. In his spare time, he has been a senior contributing editor for several local, national, and international magazines.

Craig Wright has personally conducted in excess of 1,200 IT security-related engagements for more than 120 Australian and international organizations in the private and government sectors and now works for BDO Kendall's in Australia.

In addition to his consulting engagements, Craig has also authored numerous IT security-related articles. He also has been involved with designing the architecture for the world's first online casino (Lasseter's Online) in the Northern Territory. He has designed and managed the implementation of many of the systems that protected the Australian Stock Exchange. He also developed and implemented the security policies and procedural practices within Mahindra and Mahindra, India's largest vehicle manufacturer.

He holds (among others) the following industry certifications: CISSP (ISSAP & ISSMP), CISA, CISM, CCE, GNSA, G7799, GWAS, GCFA, GLEG, GSEC, GREM, GPCI, MCSE, and GSPA. He has completed numerous degrees in a variety of fields and is currently completing both a master's degree in statistics (at Newcastle) and a master's degree in law (LLM) specializing in international commercial law (E-commerce Law). Craig is planning to start his second doctorate, a PhD in economics and law in the digital age, in early 2008.

Contents

Chapter 1 Computer Forensics in Today's World 1
Introduction 2
History of Forensics 3
Objectives of Computer Forensics 3
Computer-Facilitated Crimes 4
Reasons for Cyber Attacks. 6
Computer Forensic Flaws and Risks. 6
Modes of Attack 6
Stages of Forensic Investigation in Tracking Computer Crime. 6
Rules of Computer Forensics. 7
Digital Forensics 7
Assessing the Case: Detecting/Identifying the Event/Crime. 8
Preservation of Evidence: Chain of Custody 8
Collection: Data Recovery, Evidence Collection 8
Examination: Tracing, Filtering, Extracting Hidden Data 10
Analysis. 10
Approach the Crime Scene 11
Where and When Do You Use Computer Forensics? 12
Legal Issues 12
The Computer Forensics Lab. 13
Laboratory Strategic Planning for Business 14
Philosophy of Operation 14
A Forensics Laboratory Is a Business Venue. 14
A Forensics Laboratory Is a Technology Venue 15
A Forensics Laboratory Is a Scientific Venue 15
A Forensics Laboratory Is an Artistic Venue 15
Core Mission and Services 15
Revenue Definition. 16
"I Know How Expensive I Am. Now, How Do I Get Paid?" 16
SOP (Standard Operating Procedure) 17
Quality Standards: Accreditation. 17
Quality Standards: Auditing 18
Human Talent 18
Education and Continuing Education 19

Elements of Facilities Build-out 19
Space Planning Considerations 20
Examination Environment 20
Evidence Storage 20
Network Facilities 21
Fire Protection/Suppression 21
Water Dispersion Systems 22
Gaseous Suppression 23
Chemical Suppression 24
Electrical and Power Plant Considerations 24
LAN/WAN Planning 26
HVAC 27
Abatements 27
Static Electricity 28
EMI (electromagnetic interference) 28
Acoustic Balancing 28
Security 29
Evidence Locker Security 29
General Ambience 30
Spatial Ergonomics 30
A Note on "common office technology" 31
Personal Workspace Design 31
Common Area Considerations 31
Essential Laboratory Tools 32
Write Blockers 33
Write Block Field Kits 34
Hardware Duplication Platforms 35
Portable Forensics Systems 36
Portable Enterprise Systems 38
Laboratory Forensics Systems 38
Media Sterilization Systems 40
Data Management (Backup, Retention, Preservation) 41
CD/DVD Hardware Solutions 42
Portable Device Forensics, Some Basic Tools 43
Faraday Devices as Applied to Forensics 43
Real-World Examples 44
Portable Devices and Data Storage 45
Locating the Data 45
Power 45

Readers, readers, readers! 45
Cables, cables, cables! 45
Forensic Software 46
Operating Systems 47
File Systems 47
Investigative Platforms 48
Other/Specialty Tools 49
Tools in the Enterprise 49
Ad Hoc scripts and programs 50
Software Licensing 50
Tool Validation 50

Chapter 2 Digital Forensics: An Overview 53

Introduction 54
Digital Forensic Principles 54
Practice Safe Forensics 54
Establish and Maintain a Chain of Custody 55
Minimize Interaction with Original Evidence 57
Use Proven Tools and Know How They Work 59
Is the Tool in General Use? 60
What Is the History of the Developer and the Tool? 60
Do You Know How the Tool Works? 61
Conduct Objective Analysis and Reporting 62
Digital Environments 63
Corporate 63
Government 63
Academic 64
The Internet 64
The Home 64
Digital Forensic Methodologies 64
Litigation Support 65
Identification 65
Collection 66
Organization 66
Presentation 67
Digital Media Analysis 67
Identification 68
Collection 69
Analysis 71

Network Investigations 73
Identification 74
Collection 75
Analysis 77
Summary 78
Solutions Fast Track. 78
Frequently Asked Questions. 80

Chapter 3 Developing an Enterprise Digital Investigative/Electronic Discovery Capability 83
Introduction 84
Identifying Requirements for an Enterprise Digital Investigative/Electronic Discovery Capability 85
Costs. 87
Time 88
Resources 89
Allies. 89
Administrative Considerations for an Enterprise Digital Investigative/Electronic Discovery Capability 90
Policy and Standard Operating Procedures 91
Funding 101
Organizational Framework. 102
Training 103
Tool Validation. 103
Certification 104
Accreditation. 105
Identifying Resources (Software/Hardware/Facility) for Your Team 105
Software 106
Hardware and Storage 106
Hardware 107
Storage 107
Write Blockers 108
Facility 108
Location 108
Security 110
Ventilation and Air-Conditioning Systems. 110
Electrical and Power Systems 110
Summary 112
References 112
Frequently Asked Questions. 113

Chapter 4 Integrating a Quality Assurance Program in a Digital Forensic Laboratory 115
Introduction 116
Quality Planning, Quality Reviews, and Continuous Quality Improvement. 117
Deficiencies and Driving Out Error 117
Meeting Client Stated and Implied Needs 120
Continuous Quality Improvement 122
Laboratory Planning 123
The Structure of an Organization's SOPs or QAMs 125
"Do" or Executing the Plan 128
"Check" or Study Processes. 131
"Act" or Adapt and Refine the Plan 132
Continuous Upward Spiral of Excellence 133
Cost of Quality: Why Bother? 133
Other Challenges: Ownership, Responsibility and Authority 135
Management's Responsibility for Ownership in the Quality System 136
The Quality Manager 137
Personalities and Patience. 139
Assess Your Client's Needs 141
Adapt to Your Client's Needs 142
Private Sector Challenge 142
Summary 144
Frequently Asked Questions. 145

Chapter 5 Balancing E-discovery Challenges with Legal and IT Requirements 147
Introduction 148
Drivers of E-discovery Engineering 148
Storage 149
Federal Rules of Civil Procedure 150
Purpose 150
Costs. 151
Locations, Forms and Preservation of Electronically Stored Information. 152
Locations of ESI 153
Forms of ESI. 154
File Types 154
Metadata Fields. 155

Legal and IT Team Considerations for Electronic Discovery 156
IT Members within the Legal Team 157
Records and Information Managers 158
Information Lifecycle Managers. 159
E-mail, IM, and PDA Managers 159
Backup and Archiving Managers 161
Are You Litigation Ready? 162
Served with a Request. 162
Contact Your Chief Information Officer or Equivalent 162
Be Prepared to Field Questions from the Professionals 162
Be Prepared to Ask Questions 163
Interviews. 163
Inventory 164
Discovery Readiness Planning 165
Project Scope/Collect Available Information 166
Interviews. 166
Data Cataloging/Mapping. 168
Review of Information Collected 170
Gap Analysis 170
Findings and Recommendations 170
Business Process Improvement 172
E-discovery Tools 172
Summary 174
Frequently Asked Questions. 175

Chapter 6 Forensic Software and Hardware. 177

Introduction 178
Part 1: Forensic Software Tools. 178
Visual TimeAnalyzer 178
X-Ways Forensics 179
Evidor. 181
Slack Space & Data Recovery Tools 181
Ontrack 181
DriveSpy 182
Data Recovery Tools 182
Device Seizure 182
Forensic Sorter 183
Directory Snoop 184
Permanent Deletion of Files. 184
PDWipe 184
Darik's Boot and Nuke (DBAN) 185

File Integrity Checker 185
FileMon 185
File Date Time Extractor (FDTE) 186
Decode - Forensic Date/Time Decoder 186
Disk Imaging Tools 186
Snapback DatArrest 187
Partition Managers: Partimage 187
Linux/UNIX Tools: Ltools and Mtools 187
LTools 187
MTools 188
The Coroner's Toolkit (TCT) and Tctutils 188
Password Recovery Tools 188
@Stake 188
Decryption Collection Enterprise 189
AIM Password Decoder 189
MS Access Database Password Decoder 189
FavURLView - Favorite Viewer 189
NetAnalysis 190
Multipurpose Tools 190
Maresware 190
LC Technologies Software 191
WinHex Specialist Edition 191
Prodiscover DFT 191
Toolkits 192
NTI Tools 192
R-Studio 193
Datalifter 193
Forensic Toolkit (FTK) 194
Image Master Solo and Fastbloc 194
Encase 195
E-mail Recovery Tools 196
Network E-mail Examiner 197
Oxygen Phone Manager 197
SIM Card Seizure 198
Autoruns 199
HashDig 200
Patchit 200
PowerGREP 200
Reverse Engineering Compiler 200

Part 2: Forensic Hardware Tools 233
Hard Disk Write Protection Tools. 233
NoWrite 233
FireWire DriveDock 234
LockDown 234
Write Protect Card Reader 235
Drive Lock IDE 235
Serial-ATA DriveLock Kit. 235
Wipe MASSter 236
ImageMASSter Solo-3 IT 236
ImageMASSter 4002i 237
ImageMASSter 3002SCSI 238
Image MASSter 3004 SATA 239
Summary 241
Frequently Asked Questions. 242

Chapter 7 Incident Response: Live Forensics and Investigations 243
Introduction 244
Postmortmem versus Live Forensics 244
Evolution of the Enterprise 245
Evolution of Storage 245
Encrypted File Systems 247
Today's Live Methods 255
Case Study: Live vs. Postmortem 256
Computer Analysis for the Hacker Defender Program. 260
Network Analysis 261
Summary 262
Special Thanks 262
References 262
Frequently Asked Questions. 263

Chapter 8 Seizure of Digital Information 265
Introduction 266
Defining Digital Evidence 269
Digital Evidence Seizure Methodology 272
Seizure Methodology in Depth 274
Step 1: Digital Media Identification 275
Step 2: Minimizing the Crime Scene by Prioritizing the Physical Media 275

Step 3: Seizure of Storage Devices and Media. 276
To Pull the Plug or Not to Pull the Plug, that Is the Question 277
Factors Limiting the Wholesale Seizure of Hardware 278
Factors Limiting Wholesale Seizure: Size of Media 279
Factors Limiting Wholesale Seizure: Disk Encryption 279
Factors Limiting Wholesale Seizure: Privacy Concerns. 280
Factors Limiting Wholesale Seizure: Delays Related to Laboratory Analysis 281
Protecting the Time of the Most Highly Trained Personnel 282
The Concept of the First Responder 284
Other Options for Seizing Digital Evidence 286
Responding to a Victim of a Crime Where Digital Evidence Is Involved 288
Seizure Example 289
Previewing Information On-scene to Determine the Presence and Location of Evidentiary Data Objects 291
Obtaining Information from a Running Computer. 292
Imaging Information On-Scene 294
Imaging Finite Data Objects On-Scene 295
Use of Tools for Digital Evidence Collection 297
Common Threads within Digital Evidence Seizure 299
Determining the Most Appropriate Seizure Method 302
Summary 304
Works Cited 306
Additional Relevant Resources 308
Frequently Asked Questions. 309

Chapter 9 Conducting Cyber Investigations. 311

Introduction 312
Demystifying Computer/Cyber Crime 312
Understanding IP Addresses 315
The Explosion of Networking 318
Hostname 320
MAC Address 320
The Explosion of Wireless Networks 322
Hotspots 322
Wardriving 323
Wireless Storage Devices 325
Interpersonal Communication 325
E-mail. 326
Chat/Instant Messaging 327

Social Networking and Blogging 327
Media and Storage 328
Summary 329
Frequently Asked Questions 330

Chapter 10 Acquiring Data, Duplicating Data, and Recovering Deleted Files 333
Introduction 334
Recovering Deleted Files and Deleted Partitions 334
Deleting Files 335
Command Line Delete 335
Moving Files 337
Disk Cleanup 338
Permanently Destroying Data 339
Recycle Bin 340
What Gets Deleted 341
Configuring the Recycle Bin 341
Storage Locations of the Recycle Bin 343
Undeleting or Permanently Deleting a File 344
Damaged Recycle Bins 345
Data Recovery in Linux 347
Recovering Deleted Files 348
Deleted File Recovery Tools 349
Undelete Tools 349
Recycle Bin Replacements 358
CD/DVD Data Recovery 359
Microsoft Office Repair and Recovery 360
Compressed Files 361
Deleted Images 362
Recovering Deleted Partitions 364
Deleting Partitions Using Windows 364
Deleting Partitions from the Command Line 365
Deleted Partition Recovery Tools 369
Active@ Partition Recovery 370
Active@ Disk Image 371
DiskInternals Partition Recovery 371
GetDataBack 372
NTFS Deleted Partition Recovery 372
Handy Recovery 372
Acronis Recovery Expert 373

TestDisk 373
Scaven 373
Recover It All! 373
Partition Table Doctor 373
Data Acquisition and Duplication 374
Data Acquisition Tools 377
FTK Imager 377
SafeBack 379
DriveSpy 379
Mount Image Pro 380
DriveLook 380
DiskExplorer 381
SnapBack DatArrest 383
SCSIPAK 383
IBM DFSMSdss 383
Hardware Tools 383
ImageMASSter Solo-3 Forensic 384
LinkMASSter-2 Forensic 384
ImageMASSter 6007SAS 384
RoadMASSter-3 384
Disk Jockey IT 385
Backing Up and Duplicating Data 385
Acquiring Data in Linux 386
DD 386
Netcat 388
Summary 391
Frequently Asked Questions 392

Chapter 11 Forensic Discovery and Analysis Using BackTrack 393

Introduction 394
Digital Forensics 396
Acquiring Images 396
Linux dd 398
Linux dcfldd 408
dd_rescue 411
Forensic Analysis 413
Autopsy 414
mboxgrep 417
memfetch 419
Memfetch Find 422

pasco 424
Rootkit Hunter 426
The Sleuth Kit 428
The Sleuth Kit Continued: Allin1 for The Sleuth Kit 433
Vinetto 437
File Carving 441
Foremost 443
Magicrescue 445
Case Studies: Digital Forensics with the BackTrack Distribution 448
Summary 459

Chapter 12 Windows, Linux, and Macintosh Boot Processes 461
Introduction 462
The Boot Process 462
System Startup 462
POST: Power On Self Test 463
The Master Boot Record 464
Loading MSDOS 466
Loading Windows XP 467
Loading Linux 468
LILO Booting 469
GRUB Booting 469
The Macintosh Boot Process 469
EFI and BIOS: Similar but Different 470
DARWIN 470
Macintosh Forensic Software 471
BlackBag Forensic Suite 472
Directory Scan 473
FileSpy 473
HeaderBuilder 474
Other Tools 475
Carbon Copy Cloner 476
MacDrive6/7 478
Summary 481
Frequently Asked Questions 482

Chapter 13 Windows and Linux Forensics 483
Introduction 484
Windows Forensics 484

Where Can You Locate and Gather Evidence on a Windows Host? 484
How Can You Gather Volatile Evidence? 485
The Features and Advantages of Windows Forensics Tools 486
What Is File Slack? How Can You Investigate Windows File Slack? 495
How Do You Examine File Systems? 496
Built-in Tool: Sigverif 496
The Word Extractor Forensic Tool 497
How Can You Interpret the Windows Registry and Memory Dump Information? 497
HKEY_LOCAL_MACHINE 497
Summary of the Features and Importance of Memory Dump 499
What Is Virtual Memory? 499
System Scanner 500
Integrated Windows Forensics Software: X-Ways Forensics and its Features 501
How Can You Investigate Internet Traces? 503
Traces Viewer 503
IECookiesView 503
IE History Viewer 503
Cache Monitor 504
How Do You Investigate System State Backups? 504
Investigating ADS Streams 505
Creating a CD-ROM Bootable for Windows XP 506
Linux Forensics 507
Why Use Linux as a Forensic Tool? 508
File System Description 508
The Primary Linux Directories 509
Mount Command 512
The Linux Boot Sequence 513
The Challenges in Disk Forensics with Linux 514
Popular Linux Forensics Tools 515
The Sleuth Kit 515
Autopsy 518
SMART for Linux 520
Penguin Sleuth 520
Forensix 522
Maresware 522

Captain Nemo 523
The Farmer's Boot CD 524
Summary 526
Frequently Asked Questions 527

Chapter 14 Investigating Network Traffic and Investigating Logs 529
Introduction 530
Overview of the OSI Model 530
Layers of the OSI Model 530
The Physical Layer 530
The Data Link Layer 530
The Network Layer 531
The Transport Layer 531
The Session Layer 531
The Presentation Layer 531
The Application Layer 531
Network Addresses and NAT 531
Network Information-Gathering Tools 532
Sniffers 533
Intrusion Detection 533
Snort 534
Gathering Snort Logs 534
Building an Alerts Detail Report 536
Alerts by IP Address 536
Building an Alerts Overview Report 539
Monitoring User Activity 541
Tracking Authentication Failures 542
Listing Failed Logons 542
Identifying Single versus Multiple Failed Logons 543
Identifying Brute Force Attacks 544
Identifying a Brute Force Authentication Attack 545
Tracking Security Policy Violations 547
Determining Logon/Logoff Behavior 547
Auditing Successful and Unsuccessful File Access Attempts 549
Auditing Unsuccessful File Access Attempts 549
Auditing Successful File Access Attempts 550
Summary 551
Frequently Asked Questions 552

Chapter 15 Router Forensics and Network Forensics 553
Introduction 554
Network Forensics 554
The Hacking Process 554
The Intrusion Process 554
Searching for Evidence 555
An Overview of Routers 556
What Is a Router? 556
The Function of a Router 556
The Role of a Router 556
Routing Tables 556
Router Architecture 557
Routing Protocols 558
RIP 558
OSPF 558
Hacking Routers 559
Router Attacks 559
Router Attack Topology 559
Denial-of-Service Attacks 560
Routing Table Poisoning 561
Hit-and-Run Attacks and Persistent Attacks 561
Investigating Routers 562
Chain of Custody 562
Volatility of Evidence 562
Case Reports 563
Incident Response 564
Compromises 564
Summary 565
Frequently Asked Questions 566

Chapter 16 Investigating Wireless Attacks 569
Introduction 570
Basics of Wireless 571
Advantages of a Wireless Network 572
Disadvantages of a Wireless Network 572
Association of Wireless AP and a Device 573
Access Control 573
Encryption 573
MAC Filtering 576
Cloaking the SSID 576

Wireless Penetration Testing 578
Search Warrants 578
Direct Connections to Wireless Access Point 579
Scanning for Wireless Access Points with Nmap 579
Scanning for Wireless Access Points with Nessus 580
Rogue Access Points 580
Wireless Connect to a Wireless Access Point 581
Information Gathering 582
Injection 584
Cracking 585
Passive and Active Sniffing 586
Logging 587
Summary 588
Frequently Asked Questions 589

Chapter 17 E-mail Forensics 591
Introduction 592
Where to Start? 592
E-mail Terminology 592
Here is an Example HELO Exchange from Wikipedia: 593
Functions of E-mail 594
Archive Types 594
Server Storage Archives 594
Lotus Notes 595
Novell GroupWise 596
Local Level Archives 596
Ingredients of E-mail 597
Mailbox Archive 598
Other Associated Files of the Archive 598
Message 599
Attachments 600
Forensic Acquisition 600
Processing Local Mail Archives 601
Step 1-Acquisition Outlook PST file 601
Step 2-Processing 602
Using Paraben's E-mail Examiner 602
Using MS Outlook for processing Outlook Express files 605
Processing Server Level Archives 606
Step 1 Acquisition 607
Step 2 Processing 607

Using OnTrack PowerControls 607
Using Paraben's Network E-mail Examiner (NEMX) 610
Deleted E-mail Recovery 613
Eudora Mail 613
Outlook PST 613
Network Archives 613

Chapter 18 Steganography and Application Password Crackers 615
Introduction 616
History of Steganography 616
The Greeks 616
The Chinese 616
The Culpers 617
Civil War Rugs 617
World War I 617
World War II 617
The Vietnam War 618
Terrorists 618
The Future of Steganography 618
Classification of Steganography 618
Background Information to Image Steganography 619
Insertion 619
Substitution 619
Example 1 619
Creation 620
Six Categories of Steganography in Forensics 620
Substitution System 620
Transform Domain Techniques 620
Spread Spectrum Techniques 621
Statistical Methods 621
Distortion Techniques 621
Cover Generation Methods 621
Types of Steganography 621
Linguistic Steganography 622
Text Semagrams 622
Technical Steganography 622
Embedding Methods 622
Least Significant Bit 622
Transform Techniques 622

Spread-Spectrum Encoding 623
Perceptual Masking 623
Application of Steganography 624
Still Images: Pictures 624
Moving Images: Video 624
Audio Files 624
Text Files 625
Steganographic File Systems 625
Hiding in Disk Space 625
Unused Sectors 625
Hidden Partitions 626
Slack Space 626
Hiding in Network Packets 626
Issues in Information Hiding 626
Levels of Visibility 626
Robustness vs. Payload 627
File Format Dependence 627
Steg Tools 627
Snow 627
Steganos 628
Gifshuffle 628
Outguess 629
Stegomagic 629
Steganography vs. Watermarking 632
Fragile 632
Robust 633
Attacking Watermarking 633
Mosaic Attack 633
2Mosaic 633
Detecting and Attacking Steganography 633
Detection 634
Statistical Tests 634
Stegdetect 634
Stegbreak 634
Visible Noise 634
Appended Spaces and "Invisible" Characters 634
Color Palettes 635
Attacking Steganography 635
Steg-Only Attack 635
Known-Cover Attack 635

Known-Message Attack 635
Known-Stego Attack 635
Chosen-Stego Attack 636
Chosen-Message Attack 636
Disabling or Active Attacks 636
Application Password Cracking 637
Types of Password Cracking 637
Guessing 637
Dictionary 637
Brute Force 638
Syllable Attack 638
Rule-Based 638
Hybrid 638
Rainbow 638
Password-Cracking Tools 639
Cain and Abel 639
LCP 640
Ophcrack 641
John the Ripper 641
Brutus 641
Rock XP 642
Common Recommendations for Improving Passwords 642
No Dictionary Words 643
No Personal Data 643
Multiple Character Sets 643
Do Not Store Weak Hashes 643
Standard Password Advice 643
Change 643
Not in More Than One Place 643
Size 644
Creation 644
Summary 645

Chapter 19 PDA and Blackberry 647

Introduction 648
PDA Background Information 648
Components of a PDA 648
PDA Forensics 648
Investigative Methods 648
Step 1: Examination 649

Step 2: Identification 649
Step 3: Collection 649
Step 4: Documentation 650
PDA Investigative Tips 650
Device Switched On 650
Device Switched Off 651
Device in Its Cradle 651
Device Not in Its Cradle 651
Wireless Connection 651
Expansion Card in Slot 651
Expansion Sleeve Removed 651
Deploying PDA Forensic Tools 652
PDA Secure 652
PDA Seizure 652
EnCase 653
Introduction to the Blackberry 653
Operating System of the Blackberry 653
Blackberry Operation and Security 653
Wireless Security 653
Security for Stored Data 654
Forensic Examination of a Blackberry 654
Acquisition of Information Considerations 654
Device is in the "Off" State 654
Device is in the "On" State 655
Password Protected 655
Evidence Collection 655
Unit Control Functions 655
Imaging and Profiling 656
Attacking the Blackberry 656
Securing the Blackberry 656
Information Hiding in a Blackberry 656
Blackberry Signing Authority Tool 657
Summary 658
Frequently Asked Questions 659

Chapter 20 MP3 Forensics 661

Introduction 662
History 662
Why Is an iPod Considered Alternative Media? 663
Imaging and Hashing 663

Hardware vs. Nonhardware Imaging 664
Removing the Hard Drive 664
Linux 665
Registry Keys 671
Types of iPods 672
File Types Supported 672
File Systems 672
"Hacking Tools" and Encrypted Home Directories 673
Evidence: Normal vs. Not Normal 673
Uncovering What Should Not Be There 681
Analysis Tools 685
Summary 686

Index 687

Chapter 1

Computer Forensics in Today's World

Solutions in this chapter:

- **Computer Forensics**
- **History of Computer Forensics**
- **Objectives of Computer Forensics**
- **Computer Facilitated Crimes**
- **Reasons for cyber attacks**
- **Computer Forensics Flaws and risks**
- **Mode of Attacks**
- **Stages of Forensic Investigation in Tracking Cyber Criminals**
- **Rules of computer forensics**
- **Digital Forensics**
- **Approach the crime scene**
- **Where and when do you use Computer Forensics**
- **Legal Issues**
- **The Computer Forensics Lab**
- **Laboratory Strategic Planning for Business**
- **Elements of Facilities Build-out**
- **Essential Laboratory Tools**

Introduction

As is often the case with security compromises, it's not a matter of *if* your company will be compromised, but *when*.

If I had known the employee I hired was going to resign, break into my office, and damage my computers in the span of three days, hindsight being 20/20, I would have sent notification to the security guards at the front door placing them on high alert and made sure he was not granted access to the building after he resigned. Of course, I in hindsight, I should have done a better job of hiring critical personnel .He was hired as a computer security analyst and security hacker instructor; and was (or should have been) the best example of ethical conduct.

Clearly, we see only what we want to see when hiring staff and you won't know whether an employee is ethical until a compromise occurs. Even if my blinders had been off, I would have never seen this compromise coming. It boggles the mind to think that anyone would ruin or jeopardize his career in computer security for so little. But he did break into the building and he did damage our computers, and therefore he will be held accountable for his actions, as detailed in the following forensic information. Pay attention when the legal issues are reviewed. You will learn bits and pieces regarding how to make your life easier by knowing what you really need to know "when" your computer security compromise occurs.

Computer forensics is the preservation, identification, extraction, interpretation, and documentation of computer evidence. In Chapter 9 of *Cyber Crime Investigations*, digital forensics is referred to as "the scientific acquisition, analysis, and preservation of data contained in electronic media whose information can be used as evidence in a court of law."[1.]

In the case involving the Hewlett-Packard board of directors, seasoned investigators within HP and the primary subcontracting company sought clarity on an investigative method they were implementing for an investigation. The investigators asked legal counsel to determine whether the technique being used was legal or illegal. Legal counsel determined that the technique fell within a gray area, and did not constitute an illegal act. As a result, the investigators used it and were later arrested. This situation could befall any cyber crimes investigator.

In the Hewlett-Packard case, legal counsel did not fully understand the laws relating to such methodologies and technological issues. The lesson for investigators here is not to assume that an action you've taken is legal just because corporate counsel told you it was. This is especially true within the corporate arena. In the HP case, several investigators were arrested, including legal counsel, for their actions.

This chapter will review computer security today, the history of computer forensics, and its objectives. It will also discuss computer-facilitated crimes and the reasons for cyber crime, as well as computer forensics flaws and risks, modes of attack, digital forensics, and the stages of forensic investigation in tracking cyber criminals.

History of Forensics

Forensics has been around since the dawn of justice. Cavemen had justice in rules set to protect home and hearth. Francis Galton (1822–1911) made the first recorded study of fingerprints, Leone Lattes (1887–1954) discovered blood groupings (A, B, AB, and 0), Calvin Goddard (1891–1955) allowed firearms and bullet comparison for solving many pending court cases, Albert Osborn (1858–1946) developed essential features of document examination, Hans Gross (1847–1915) made use of scientific study to head criminal investigations. And in 1932, the FBI set up a lab to provide forensic services to all field agents and other law authorities across the country. When you look back at these historic forensic events, we see patterns of confidence in the forensic information recovered and analyzed. You will see in book, today's computer forensics is clearly a new pattern of confidence, acceptance and analysis.

Objectives of Computer Forensics

Cyber activity has become an important part of the everyday lives of the general public. According to the EC Council, eighty-five percent of businesses and government agencies have detected a security breach. The examination of digital evidence (media) has provided a medium for forensic investigators to focus on after an incident has occurred. The ultimate goal of a computer forensic investigator is to determine the nature and events concerning a crime and to locate the perpetrator by following a structured investigative procedure.

> What is forensic computing? A methodical series of techniques and procedures for gathering evidence, from computing equipment and various storage devices and digital media, that can be presented in a court of law in a coherent and meaningful format.
>
> —Dr. H.B. Wolfe

Investigators must apply the following two tests for evidence for both computer forensics and physical forensics to survive in a court of law:

Authenticity Where does the evidence come from?

Reliability Is the evidence reliable and free of flaws?

With that said, when determining whether you want to conduct a computer crime investigation, that should be predetermined through policy and what is "acceptable risk" to your company.

Cyber crime includes the following:

- **Theft of intellectual property** This pertains to any act that allows access to patent, trade secrets, customer data, sales trends, and any confidential information.

- **Damage of company service networks** This can occur if someone plants a Trojan horse, conducts a denial of service attack, installs an unauthorized modem, or installs a back door to allow others to gain access to the network or system.
- **Financial fraud** This pertains to anything that uses fraudulent solicitation to prospective victims to conduct fraudulent transactions.
- **Hacker system penetrations** These occur via the use of sniffers, rootkits, and other tools that take advantage of vulnerabilities of systems or software.
- **Distribution and execution of viruses and worms** These are some of the most common forms of cyber crime.

Cyber crime comprises the "3 Ts": tools to commit the crime, targets of the crime (victim), and material that is tangential to the crime.

Cyber crime is motivated by many different things. Often it's the thrill of the chase, and a desire for script kiddies to learn. Sometimes cyber crime is committed by psychologically motivated criminals who need to leave a mark, or simple misguided trust in other individuals that they are not leading you astray. Other times such crimes are committed by a person or group that is out for revenge; perhaps it's a thwarted employee or friend that wants to embarrass the target. Most likely, a cyber criminal is being paid to gain information; hackers involved in corporate espionage are the hardest to uncover and often are never seen.

Damage & Defense…

Curbing Computer Crime

Computer crime happens more often than car accidents, and car accidents occur four times a minute in the United States. A defensive posture, security awareness training, and continuous good communication help keep insider threats to a manageable minimum.

Computer-Facilitated Crimes

Our dependency on the computer has given way to new criminal opportunities. Computers are increasingly being used as a tool for committing crimes, and as such they are posing new challenges for investigators, for the following reasons:

- The proliferation of PCs and Internet access has made the exchange of information quick and inexpensive.

- The use of easily available hacking tools and the proliferation of underground hacking groups have made it easier to commit cyber crimes.
- Anonymity allows anyone to hide his identity while committing crimes.
- E-mail spoofing, creating fake profiles, and committing identity theft are common occurrences, and there is nothing to stop it, making investigation difficult.
- With cyber crimes, there is no collateral or forensic evidence, such as eye witnesses, fingerprints, or DNA, making these crimes much harder to prosecute.

Damage & Defense…

Bridging the Gaps

Real-Life Solutions: One of my first cases involved a woman whose ex-boyfriend was impersonating her online. He created an online user profile using her personal information and her picture on a popular chat site. During his chats, while pretending to be her, he solicited sexual acts from several men and gave her contact information to them. This information included her home address. During several of these online chats, he described a rape fantasy she wanted to fulfill with the men he was chatting with. When discussing the case with the prosecutor's office, we brainstormed about the charges we would use. There were no identity theft laws in place at that time. So, we decided to use traditional charges, including reckless endangerment, aggravated harassment, and impersonation. Here is an outline of our justification for using these statutes:

- We selected reckless endangerment because the men were visiting the victim's home expecting to engage in sexual acts with her. These acts included the rape fantasy that the suspect described during the online chats. The reckless endangerment aspect of this crime was the possibility of some male raping her because of the described rape fantasy the suspect spoke about. Someone could have really raped her.
- We selected aggravated harassment because of the number of phone calls she was receiving day and night that were sexually explicit. In New York, it covered the annoying phone calls the victim was getting.
- We chose the charge of impersonation because the ex-boyfriend was pretending to be her. This impersonation included more than him just pretending to be her online. It included giving out all of her personal information, along with her picture. Today, this would most probably be covered under an identity theft law.

Reasons for Cyber Attacks

Today, cyber attacks are committed by individuals who are more organized. Cyber crime has different connotations depending on the situation. Most of us equate cyber crime with what we see on TV and in the news: porn, hackers gaining access to sensitive government information, identity theft, stolen passwords, and so on. In reality, these types of computer crimes include more often than not, theft of intellectual property, damage of company service networks, embezzlement, copyright piracy (software, movie, sound recording), child pornography, planting of viruses and worms, password trafficking, e-mail bombing, and spam.

Cyber criminals are taught to be more technically advanced than the agencies that plan to thwart them. And today's criminals are more persistent than ever. According to the EC Council, *computer crime is any illegal act involving a computer, its system, or its applications*. A computer crime is intentional, not accidental (we discuss this in more detail in the "Legal Issues" section, later in this chapter).

Computer Forensic Flaws and Risks

Computer forensics is in its developmental stage. It differs from other forensic sciences as digital evidence is examined. There is a little theoretical knowledge to base assumptions for analysis and standard empirical hypothesis testing when carried out lacks proper training or standardization of tools, and lastly it is still more 'art" than "science.

Modes of Attack

There are two categories of cyber crime, differentiated in terms of how the attack takes place:

- **Insider attacks** These involve a breach of trust from employees within an organization.
- **External attacks** These involve hackers hired by either an insider or an external entity whose aim is to destroy a competitor's reputation.

Stages of Forensic Investigation in Tracking Computer Crime

A computer forensic investigator follows certain stages and procedures when working on a case. First he identifies the crime, along with the computer and other tools used to commit the crime. Then he gathers evidence and builds a suitable chain of custody. The investigator must follow these procedures as thoroughly as possible. Once he recovers data, he must image, duplicate, and replicate it, and then analyze the duplicated evidence. After the

evidence has been analyzed, the investigator must act as an expert witness and present the evidence in court. The investigator becomes the tool which law enforcement uses to track and prosecute cyber criminals.

For a better understanding of the steps a forensic investigator typically follows, consider the following, which would occur after an incident in which a server is compromised:

1. Company personnel call the corporate lawyer for legal advice.
2. The forensic investigator prepares a First Response of Procedures (FRP).
3. The forensic investigator seizes the evidence at the crime scene and transports it to the forensic lab.
4. The forensic investigator prepares bit-stream images of the files and creates an MD5 # of the files.
5. The forensic investigator examines the evidence for proof of a crime, and prepares an investigative report before concluding the investigation.
6. The forensic investigator hands the sensitive report information to the client, who reviews it to see whether they want to press charges.
7. The FI destroys any sensitive client data.

It is very important that a forensic investigator follows all of these steps and that the process contains no misinformation that could ruin his reputation or the reputation of an organization.

Rules of Computer Forensics

A good forensic investigator should always follow these rules:

- Examine original evidence as little as possible. Instead, examine the duplicate evidence.
- Follow the rules of evidence and do not tamper with the evidence.
- Always prepare a chain of custody, and handle evidence with care.
- Never exceed the knowledge base of the FI.
- Make sure to document any changes in evidence.

If you stay within these parameters your case should be valuable and defensible.

Digital Forensics

Digital forensics includes preserving, collecting, confirming, identifying, analyzing, recording, and presenting crime scene information.

Assessing the Case: Detecting/ Identifying the Event/Crime

In any type of investigation, the computer forensic examiner must follow an investigation process. That process begins with the step of assessing the case, asking people questions, and documenting the results in an effort to identify the crime and the location of the evidence. Computer investigations are conducted on two types of computers: the computer used to commit a crime, and computer that is the target of the crime.

Preservation of Evidence: Chain of Custody

Preserving the chain of custody is the next step. Identification of the evidence must be preserved to maintain its integrity. A chain of evidence must be prepared to know who handled the evidence, and every step taken by the forensic investigator must be documented for inclusion in the final report. Sometimes a computer and its related evidence can determine the chain of events leading to a crime for the investigator as well as provide the evidence which can lead to conviction.

NOTE

A chain of custody is the accurate documentation of the movement and possession of a piece of evidence, from the time it is taken into custody until it is delivered to the court. This documentation helps prevent allegations of evidence tampering. It also proves that the evidence was stored in a legally accepted location, and it documents who is in custody and control of the evidence during the forensic testing phase.

A bit-stream image is an exact duplicate of a computer's hard drive in which the drive is copied from one drive to another, bit by bit. This image is then authenticated to the original by matching a digital signature, which is produced by a mathematical algorithm (usually the MD5 standard) to ensure that no changes have occurred. This method has become the de facto standard and is widely accepted by the industry and the legal system.

Collection: Data Recovery, Evidence Collection

Finding the evidence, discovering relevant data, preparing an Order of Volatility, eradicating external avenues of alteration, gathering the evidence, and preparing a chain of custody are the recommended processes for collecting data. After you collect data, you should create an MD5 hash of the evidence. Prior to collection, one should do preliminary assessment to search for the evidence. After the assessment is concluded, collect and seize the equipment used in committing the crime, document the items collected, such as floppy disks, thumb

drives, CDs, DVDs, and external back up drives. A photo of the crime scene should be taken before removing the evidence.

Damage & Defense…

Hashes

Hashes use cryptographic algorithms to create a message digest of the data and represent it as a relatively small piece of data. The hash can be used to compare a hash of the original data to the forensic copy. When the hashes match, it is accepted as proof that the data is an exact copy. Although it has not been challenged yet, the traditional hashes of CRC, MD5, and SHA1 have been cracked. Also, there are limitations in the sheer volume of 128 bit hashing algorithms such as MD5. There are only 2^{128} possible MD5 hashes. If the large multi-terabyte file server being analyzed stores $2^{128} + 1$ files, there absolutely will be two different files with unique data with the same hash. Now it is understood that 2^{128} is about 340 billion, and it would be an extremely large storage array of tiny files, but this fact opens the door for doubt, which could ruin a criminal prosecution. Although 2^{128} is still a huge number, as storage grows, it is not unrealistic to believe that 128 bit hashes will become an increasing issue. It will probably be an issue on large storage systems long before it becomes as big an issue on single workstations. The future appears to be the use of the SHA-256 algorithm and other 256 bit hashes. For now, the National Software Reference Library Hashes use the SHA-1 and MD5 algorithms.

After collecting all the information, the investigator can then list the steps that can be taken during the investigation and then begin. Caution, it is not necessary to seize the entire system. Identify the relevant data and copy that, otherwise it can result in over collection.

Notes from the Underground…

Suggested Tool Kit Contents

Your tool kit should contain the following components:

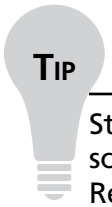
Hardware Target hard drives, write blocker, and cables (network, IDE, and SCSI)

Software Boot disks and drivers for both your forensic system and any system you may encounter, especially for network cards

Continued

Tools Allen keys; large and small screwdrivers (standard, Phillips, and Torx)

Other content Labels , anti-static bags, pens and markers, blank media: (CDs, DVDs), and a camera



TIP

Sterilize all the media to be used in the examination process, enter the crime scene, take a snap shot of the scene and then carefully scan the data sources, Retain and document the state and integrity of items at the crime scene then transport the evidence to the forensic facility

Examination: Tracing, Filtering, Extracting Hidden Data

The examination process follows the collection process. The computer forensic investigator must trace, filter, and extract hidden data during the process. Some evidence cannot stay for long. Such evidence is called volatile evidence because it needs consistent power supply for storage. There is also evidence that contains the information that keeps changing. Investigators must review registers and cache, routing tables, ARP cache, process tables, and kernel statistics and modules.

Harlan Carvey looks at the Order of Volatility from a "live system" view. volatile data must be preserved in order of volatility, with the most volatile data preserved first. This applies to live systems for the most part, but the way in which we approach live systems will become more important in the near future. An example of an order of recovery of system data according to volatility looks like this:

- **Virtual memory** Swap space or paging files
- **Physical disks** The physical hard disks of a system
- **Backups** Offline back-up media such as magnetic tape or other media: It is extremely possibly the data you are looking for may not be on the system today, but it was there yesterday and is on last night's backup.

Analysis

Analysis of the data is greatly different from retrieving the evidence and depends greatly on exactly how the copy is. There are various techniques to capture an exact forensic copy to copy the evidence disk so you can analyze the data. Analysis should be done on the duplicate copy

so that the original evidence can be protected from alteration because the first rule of forensics is to preserve the original evidence. Once a copy is created, use the copy for further processes. Analysis can be carried out using various forensic analysis tools such Encase, Access Data etc.

Damage & Defense…

Digital Evidence

When digital evidence is extracted from digital resources an investigator must:

Protect the subject computer system during the forensic examination from any possible alteration, damage, data corruption, or virus introduction.

Discovers all files on the subject system.

This includes existing normal files, deleted yet remaining files, hidden files, password-protected files, and encrypted files.

Recover all (or as much as possible) of discovered **deleted files**.

Reveal (to the extent possible) **the contents of hidden files as well as temporary or swap files** used by both the application programs and the operating system.

Access (if possible and if legally appropriate) the contents of **protected or encrypted files**.

Analyze all possibly relevant data found in special areas of a disk.

This includes 'unallocated' space on a disk, 'slack' space in a file and disk cluster.

Print out an overall analysis of the subject computer system, as well as a listing of all possibly relevant files and discovered file data.

Provides an opinion of the system layout, the file structures discovered, any discovered data and authorship information, any attempts to hide, delete, protect, encrypt information, and anything else that has been discovered and appears to be relevant to the overall computer system examination.

Provide expert consultation and/or testimony, as required.

Approach the Crime Scene

Due to the presence of a majority of electronic documents, and the skills necessary to search and identify data in a computer, combined with the fact that digital evidence is delicate in nature for recovering deleted, encrypted or, corrupted files from a system there is a growing need for Forensic Investigators to approach crime scenes.

An investigator, if trained properly, will ensure that no possible evidence is damaged, destroyed, or compromised by the forensic procedures used to investigate the computer. (Preservation of evidence).

No computer malware, or harmful software, is introduced to the computer being investigated. (Non-contamination of evidence). Any extracted or relevant evidence is properly handled and protected from later mechanical or electromagnetic damage (extraction and preservation of evidence). A continuing chain of custody is established and maintained (Accountability of evidence).and that normal operations are effected for a limited amount of time.(limited interference of the crime scene on normal life).

Where and When Do You Use Computer Forensics?

When there is a need to provide Real Evidence such as reading bar codes, magnetic tapes and to identify the occurrence of electronic transactions and reconstruct an incidence with sequence of events. You use computer forensics when a breach of contract occurs, or if copyright and intellectual property theft/misuse happens or during employee disputes where there is damage to Resources.

Legal Issues

It is not always possible for a computer forensics expert to separate the legal issues surrounding the evidence from the practical aspects of computer forensics Ex: The issues related to authenticity, reliability and completeness and convincing. The approach of investigation diverges with change in technology. Evidence shown is to be untampered with and fully accounted for, from the time of collection to the time of presentation to the court. Hence, it must meet the relevant evidence laws

Damage & Defense…

Permission

When my company was broken into, I provided verbal permission to law enforcement to search my facility and locate the missing computers. I also gave permission to turn on one of the computers where we confirmed the x-employee had broken into the building, stolen the computers, accessed the computers, erased intellectual property and left the building hiding the computers.

There are legal concerns, not just technical concerns. For example, for some forensic monitoring activity a certain level of security may be legally required, or your ability to monitor certain kinds of activities may be restricted. Also, should you ever need to prosecute; your logs may not be admissible in court. Local and federal laws must be considered when devising a security policy.

The computer revolution has given way to white collar crimes done on the internet. Remote targets are compromised by malicious users daily. While investigating these crimes, International issues can be raised as the electronic evidence necessary to prevent, investigate or prosecute a crime is located outside the borders of the country and law enforcement must seek assistance from law enforcement authorities in the different country. Preservation of evidence or request for evidence can be made under mutual legal assistance agreements or if no assistance is forthcoming through the Letters Rogatory process.

Consistency with all legal systems, the ability to implant confidence in the integrity of evidence, allowances for the use of common language and applicability at ever level are confronted by investigators.

Computer law is a large field. Areas of concern to security administrators are: what constitutes illegal use of a computer, what you can and can't do to detect or monitor it, the status of any evidence you may collect, and your exposure to civil liability suits in event of a security problem. Computer crime law is a new field. The statutes are quite recent, less than 10 years old with little case law for guidance. Interpretations may change, and the laws themselves may change, as legislators react to newer threats.

The Computer Forensics Lab

The process of implementing and operating a computer forensics laboratory could be the subject of an entire series of books. This chapter, however, will attempt to share a few ideas regarding core concepts to be considered during the planning, construction and operation of a data forensics facility. The chapter's bias will be toward mid-level size operations (corporate installations and stand-alone facilities) in order to demonstrate a diversity of concepts relating to facilities planning, business operations and service offerings.

Recent changes to the Federal Rules of Civil Procedure (FRCP) in December 2006 have impacted the manner in which digital information is managed in civil litigation. The FRCP formalized the role of digital information in a legal environment. The rules have formally identified the role of Electronic Stored Information (ESI) and how it will be handled and presented in a judicial setting.

The advent of personal computing empowered individuals to create and manage information on a massive scale; the vast majority of information created now exists in digital form on some type of computing system. An entire field of data analysis and digital investigation has evolved in response to the threat of wrongdoing in this digital realm. The technology (laptops, desktops, cell phones, the Internet) empowering individual productivity and

creativity is the same technology used to conduct activity against company policy or in violation of the law. Corporate investigators and law enforcement officers need to have the capability to investigate these types of digital transactions by identifying, recovering, analyzing and reporting on the digital facts. The role of data forensics analysis will be of increasing importance to the legal system as information continues to evolve into the purely digital and the systems upon which that information is stored become more technologically advanced. The need and demand for expert forensics examiners and forensic data investigation facilities will likewise be on the rise.

Laboratory Strategic Planning for Business

The topic of strategic planning for business development is a series of books unto itself.

A few points of interest will be touched upon as of special interest in developing a forensics practice: philosophy of operation, core mission and services, revenue definition, and SOP definition.

Philosophy of Operation

Every data forensics implementation will reflect four core modes of operation. From solo-practitioner operations to government investigative arm, forensics implementations will function according to a similar set of operating philosophies. The core four aspects of operation are the business operations aspect, the technology venue aspect, the scientific practice aspect, and the artistic expression aspect. Regardless of scope, a computer forensics initiative must pursue sound business practices, must function in the realm of high technology with high technology talent as ongoing status quo and must foster excellence of method and diverse, creative vision in solving technology investigation problems.

A Forensics Laboratory Is a Business Venue

Every computer forensics laboratory is a business venture. A 1099 contract solo investigator, a commercial forensics department in the civilian litigation support space, a city/state police crime lab, a law firm's internal digital investigations group and a federal network of investigative facilities are all business venues that must behave according to the principles of sound business management, financial profitability, core service provision, etc. A police crime lab may not be pursuing profit per se, but that lab has to demonstrate value of service and ROI (return on investment) in order to remain funded or acquire annual budget allocations and new technologies to continue fighting crime. A solo practitioner must remain competitive in the marketplace he/she serves with regard to cost, service provision, and continuing education. A corporate commercial forensics venture must demonstrate profitability and maintain high standards for customer service and product quality in order to remain competitive in the marketplace. A massive entity such as the United States government's network of

nationally distributed forensics facilities and allied investigative entities must still obey the principles of good business management, seek operational excellence, and demonstrate value for service and ROI to the United States Congress and Senate in order to remain funded. *Running a data forensics laboratory means running a good business at all levels of scope.*

A Forensics Laboratory Is a Technology Venue

A data forensics facility of any size is the embodiment of front-of-the-wave mastery of data and data storage technologies in all its' various guises. Criminals often afford the newest toys and desire the most complex technologies to hide their crimes from prying eyes, so the data forensics community must always strive to master technology as fast as technology evolves. The commercial consumer marketplace is always rolling out a new wave of the newest, shiniest technologies available in order to keep up with consumer demand for progress; again, the forensics community is at the front of the line, dismantling and investigating every new gadget that hits the shelves in order to reveal its secrets.

A Forensics Laboratory Is a Scientific Venue

Understanding and implementing technology isn't sufficient, however. The practice of any branch of Forensics is a practice of Science. Examiners strive to perform their duties according to reliable, repeatable, valid, objective, consistent and accurate methodologies in order to reveal facts objectively via empirical observation, deductive reasoning and conversion of hypothesis to demonstrable proof of fact, thereby empowering the presentation of findings of value to be put forth as facts of merit in the court of law.

A Forensics Laboratory Is an Artistic Venue

The investigative process is more than a rigid set of procedures. Intuition and creativity play as great a role for the Forensic examiner as do sound methodologies. Fact finding in a wildly diverse technological realm requires a great degree of technical prowess as well as a flexible mind; forensic examiners often must be artisans of technology creation and deconstruction. Raw technology skill does not empower an investigator to understand the interaction of man and machine: intuitive awareness how the tools of technology and human nature, human thought processes, and human frailties interact allows for much of the artistry and creativity of forensic investigation to be revealed.

Core Mission and Services

Foremost in the consideration of a forensics facility design plan, decide what services the facility is to provide and the scope at which it is to provide those services. A firm grasp of the prospective laboratory's core mission and scope of service will provide guidance on every aspect of building and operating that forensic facility, touching on everything from annual budget to furniture ergonomics. Based upon scope of service, a good forensics laboratory can

reside in one room or it may require an entire building with multiple teams of specialists executing diverse tasks across multiple disciplines in each of several geographic regions. A law enforcement agency will focus upon violations of criminal statutes; a governmental agency may focus on one or more aspects of civil litigation; a commercial venture will typically define a service package and then market that package to any number of audiences.

Revenue Definition

A very applicable adage applies to a data forensics facilities' operational capability: "Anything is possible with enough money, manpower and time". As always, knowing how to effectively address the five w's (who, what, when, where, why) of a business plan will dictate the completeness of the plan from concept to execution. Implement a five year strategic plan. Plan for successful growth. Plan based upon the realities of the specific environment in which the facility will reside, and to which the facilities will respond. Implement a realistic and generous budget: justify it with a cost vs. reward argument. Define milestones to achieve and a growth track to follow. Ultimately, the budget implemented will need to fully serve the needs of the facility in both actual operation and realization of strategic vision.

Every forensics facility initiative, whether law enforcement, corporate, or for-profit, will require funds to function. Developing a strong business plan based upon costs of doing business versus profitability of work product is necessary regardless of intended audience. Every operation will need to demonstrate ROI (return on investment) in order to prove the viability of the venture.

Costs of doing business will include line item tangibles such as hard dollar outlay to build, staff, stock, operate, maintain and grow a facility. Costs will also include intangibles such as administrative overhead for policy and procedure creation, implementation and ongoing process improvement. Buffer will need to exist for known business variables such as payroll fluctuation and increasing utilities costs. Equipment requires maintenance and replacement. Etc.

Defining profitability in light of any given operational ROI will vary depending on the core service provision of the facility. A law enforcement laboratory may want to define profitability in terms of metrics addressing man hours expended and cases processed vs. convictions/pleas achieved; a non-profit or governmental agency may want to define profitability in terms of an annual impact statement on their sector of influence. Commercial ventures will certainly define profitability in terms of billable professional hours, machine time and/or line item service provision. Regardless of how profitability is qualified, profitability needs to be quantified in order to demonstrate the fitness of the venture.

"I Know How Expensive I Am. Now, How Do I Get Paid?"

A data forensics operation will either position itself as a cost center or as a revenue generator. In most law enforcement and government agency scenarios, a forensics offering will be

perceived as a cost center and will rely on departmental budget allocations, grants, etc. for funding. ROI will generally be defined by demonstrating efficiency and operational excellence. Profitability will be defined in terms of ongoing results achieved.

Corporate implementations, likely to be cost centers, may define themselves as revenue generators by creating a "billback" or cross-charge system in which profitability is determined by revenue tracking demonstrated by billable units (either "credit-for-time-served" being equated to billable hours or actual inter-departmental invoicing "billed back" to the requesting business unit.)

Commercial forensics service providers will invoice for services provided and must demonstrate a net profit margin above operating costs.

SOP (Standard Operating Procedure)

Whether applied at the strategic, daily operations or process specific level, policy and procedure implementation will ultimately be the measure of operational excellence by which the caliber of a data forensics laboratory (and the product the laboratory produces) is defined. The 10k SOP should be defined while still in the planning stages of laboratory design. The ultimate goal of any work executed in a data forensics laboratory is to send valid, objective electronic evidence into a court of law. The laboratory itself must operate according to high professional standards; the employees of the laboratory must comport themselves professionally and ethically; and the tasks executed by the employees in the investigation and handling of potential evidence must be procedurally sound. "Soundness" of process should be demonstrated by testable, repeatable procedures generating predictable results. Evidence integrity must be defensible; the first defense against spoliation attacks is a defensible process. For all of these things to occur, a robust policy for procedure implementation and oversight is necessary. Workflow management, product testing, process analysis and method execution all fall within the scope of need for SOP development.

Quality Standards: Accreditation

Demonstration of operational excellence is important to any business operation. For a forensics facility of any discipline, demonstration of operational excellence is of utmost importance and independent certification of operational excellence is greatly desired. One route taken by many businesses is ISO (International Organization for Standardization) certification. A forensics laboratory could and should pursue ISO accreditations. An organization explicit to the universe of forensics (but not limited to data forensics) is the ASCLD/LAB (American Society of Crime Laboratory Directors/LAB) certifying body. ASCLD/LAB endorses a certification track for a data forensics facility that incorporates both ISO standard 17025 and a supplemental ASCLD requirement set explicit to laboratory operations. The certification itself includes both benchmark standards for operation and ongoing oversight for maintaining accreditation status.

The ASCLD/LAB model for facility operations focuses heavily upon a number of areas deemed critical to quality laboratory performance:

- Leadership quality, hierarchy and effectiveness
- Guidelines regarding policy and procedure creation and execution
- Interoffice and official communication protocols, both vertical and horizontal
- Definition of educational standards and skills testing
- Investment in human resources via training and development
- Physical plant design (security, infrastructure, fixtures)
- Locale ergonomics (personal and shared workspace)
- Implementation of business process control systems and audit methodology
- Explicit requirements at the level of business processes specific to the realm of
- evidence handling and forensic data examination

Both the ISO 17025 and ASCLD/LAB documents are a very useful high level document to review when planning both the physical plant and the operational function of a data forensics laboratory. ASCLD/LAB-International can be contacted at www.ascld-lab.org.

Quality Standards: Auditing

Demonstration of operational excellence includes the need for multiple audit channels:

- Individual procedures must be tested for validity of method and adherence to process
- Hardware and software tools require testing to prove function
- Individual competency levels need to be performance tested
- Workflow requires audit to guarantee operational excellence
- Inventory control & chain of custody require ad hoc demonstration of 100% competency
- Overall business SOP and mid-level operating procedure require constant re-assessment

A robust audit system is required in order to achieve the level of process rigor required of any forensics facility.

Human Talent

A forensics examination environment is only as good as the talent associated with the initiative. The best hardware purchasing plan in the world won't matter if the human

element does not receive the same quality of investment. Experience gathering, knowledge sharing, continual education and a serious investment in human resource development are essential to the overall success of a data forensics laboratory.

Education and Continuing Education

Bachelor's level and Master's level degree programs exist that focus upon forensic investigation; several universities offer a Criminal Justice degree with a specialty in digital forensics. Multiple certifications exist for the forensic examiner. Certification programs demonstrate both the breadth of knowledge and the hands-on proficiency of the examiner. Maintaining certification means routine re-testing and accrual of classroom training hours on a regular basis.

Available certifications include:

- **Law enforcement:**
- ***CFCE*** Certified Forensic Computer Examiner (www.cops.org) IACIS
- **Civilian and law enforcement:**
- ***CCE*** Certified Computer Examiner (www.certified-computer-examiner.com) ISFCE
- ***GCFA*** GIAC Certified Forensic Analyst (www.sans.org)
- ***CFCI*** Certified Hacker Forensic Investigator (www.eccouncil.org)
- **Software specific:**
- ***ACE*** Access Data Forensic Tool Kit (www.accessdata.com)
- ***EnCE*** Guidance Software EnCase Certified Examiner (www.guidancesoftware.com)

Elements of Facilities Build-out

In general, addressing any element of facilities build-out includes budgeting for construction and operation, provision of service based upon normal operations, provision based upon adverse events and subsequent disaster recovery, and provision based upon a roadmap for expansion, growth and future modernizations. These topics can tailor the design of facilities elements such as electrical or HVAC provision, or they can apply to business operations and workflow on an ongoing basis. Size of implementation and budget constraint always delimit facilities complexity. Small facilities may not need to address many of the listed concepts addressed herein, but the average corporate, law enforcement or stand alone facility will likely address all of them, plus more.

Space Planning Considerations

In conceptualizing the overall layout of a forensics laboratory, attention should be given to at least four functional areas: administrative area, examination space, network facilities, and evidence storageAdministrative Area

Administrative space is comprised of office space for personnel associated with the forensics team (project management, executive staff, investigators etc.), general meeting space for internal personnel and clientele, and "privacy" or guest areas. This environment should provide adequate room for team meetings and a comfortable environment for customer facing activities. The forensics investigation team will likely spend a lion's share of their time in the examination space (often a shared environment with little "personal space"). Consideration should be given to adequate private workspace where individuals can hold confidential conversations, make telephone calls and engage in general corporate communications.

Examination Environment

Examination space is the "lab proper", i.e. all space devoted to the technical and investigative aspects of the forensics examination process. This environment is the home area for all of the technical equipment associated with the examination process and will likely be the functional area of the laboratory in which the forensics technical staff spends a vast majority of their time. Access to the examination space should be restricted to relevant personnel and traffic to and from the examination space should be logged. Provide plenty of surface area and dedicate significant square footage per investigator (a good starting metric is 100 square feet, or the measure of a $10' \times 10'$ office space). Provide significant square footage for the location of forensics equipment (both shared and individual assets).

Evidence Storage

Evidence storage is dedicated storage space existing for the sole purpose of warehousing digital evidence and other evidentiary items. The evidence storage area is the physical embodiment of chain of custody functionality. Evidence storage should be the most secure/demanding environment to access, the most rigorously controlled area for any type of entry/egress/activity, and the most physically segregated area of a forensics build-out. The "evidence locker" must be constructed to defeat forced/unauthorized entry. It should be designed such that its' contents survive environmental events. All access to this environment should be controlled with the highest rigor and restricted to key personnel, often to a single Custodian of Evidence. Multiple challenges to entry and identity should be employed. The evidence storage environment will require, in many cases, customized abatements (such as EMI shielding). A robust information management system should accompany an evidence storage environment: automated security systems should be in place challenging all accessors and logging all accesses. Inventory should be controlled via both ink-signature and automated electronic management systems.

Information management systems employed should have a robust audit methodology that guarantees completeness and accuracy of the information maintained. Any and all components of the Evidence Storage Facility should assure the "who, what, when, where, and why" of every object considered "evidence" is always known and documented.

Network Facilities

This space is the environment in which data network, security and telecommunications equipment serving the laboratory space resides. Ideally, this space should be protected from compromise to the same degree that evidence storage is protected. The physical elements of data networking and security technology warehousing, transmitting or otherwise accessing evidentiary data materials or examination process work product should be dedicated and stand alone infrastructure. This rule applies to data cabling, servers, switches, routers and any other physical element of the networked technology systems serving the forensics space. Steps should be taken to assure that any inbound or outward facing day-to-day business protocols (i.e. corporate email, telephony, internet accesses, etc.) provision across a completely separate physical network architecture.

Fire Protection/Suppression

A forensics laboratory, especially a larger facility, requires a well thought out fire protection plan. With regard to overall fire code, the local Fire Marshall can provide specifics regarding local standards and ordinances; if the laboratory is to be built out in pre-existing space, the property may have its own supplemental fire protection requirements, especially if the need to tie into existing infrastructure exists. Fires are classified based upon the material serving as fuel for the fires. The fire suppression methods employed will generally be determined via understanding cost constraints, habitation zones of personnel and the technology venue residing in the space. In many ways, the ideal fire suppression system for a forensics facility will model after data center or disaster recovery data co-location facility design plans. Of special concern are Class C fires, which involve both some flammable fuel substrate and the presence of electricity. A new facility will be presented with multiple fire protection options, and the choices made regarding fire suppression implementation can have cost, timeline and design impact on every other aspect of the build-out.

Fire classification varies worldwide with regard to accepted "classes" of fire. In the United States, fire ratings fall into five main classifications.

- Class A Common (solid) combustibles.
- Class B Liquids and gases.
- Class C Fires involving electricity.
- Class D Combustible metals.
- Class K Cooking fluids/oils.

In the forensics laboratory environment, the most common fire classes are likely to be Class A (infrastructure materials) and Class C (electrical fires involving powered-up technology). In order to protect against a Class A/C hazard, multiple options are available regarding suppression system.

- Water Dispersion Systems (Air Pressurized Water Systems)
  - Wet pipe system
  - Dry pipe system
  - Preaction system
- Gaseous Suppression (Clean Agents)
  - Inert Gas
  - Flourine Compound
- Chemical Suppression
  - Foam
  - Dry Chemicals

Water Dispersion Systems

The three most common water dispersion system designs are "wet pipe", "dry pipe" and "preaction".

Wet Pipe System

This system employs a piping system that maintains a constant water load. This system is generally the most cost effective and low maintenance of all fire protection options, but does have drawbacks in an environment where significant electronics and high technology reside. Inadvertent failure or impact damage means water leaks (small or large). Typically, wet pipe systems are easy to repair and maintain, and they have a fast recovery window after activation.

Dry Pipe System

This system employs a piping system that maintains a pressurized air load. The pressurized air holds back liquid flow under normal circumstances. This system relies upon deployment (sprinkler) head events to trigger gas release, which then allows water to flow into the pipes as the gas bleeds out. Typically, dry pipes are significantly more expensive than wet pipe systems, taking more hardware to deploy, having a higher facilities space requirement (for the gas storage and pump equipment), and offering the same ultimate drawbacks as wet pipe. Additionally, dry pipe offers maintenance complexities and higher maintenance costs. Dry pipe does offer protection from pipe bursting in cold environments.

Preaction System

Preaction systems are typically the second level of fire protection implementation to be considered in a facilities build-out. This system is a modified dry pipe arrangement; the advantage of a preaction system is the use of two triggers to release the liquid suppressant. A valve, typically an electronic valve, acts as the release inhibitor; water is not held back by gas pressurization. The valve will be controlled by a discrete fire sensor (i.e. one that operates independently of any sprinkler heads, etc.). If the valve releases, the pipes fill with liquid and the system then behaves like wet pipe. A second event must occur at the level of the delivery heads in order to release water into the environment. Pipe impact damage and head failures offer less threat to the surrounding environment given the fact that the pipes are in a no-load state under normal circumstances. The potential time delay between valve sensor engagement and sprinkler engagement could also benefit the environment presuming some intervention is able to resolve a sensor-perceived threat prior to head discharge. The cost factor step from wet pipe to preaction pipe can be a significant increase as size of planned facility increases. Preaction systems have the increased complexity level and maintenance disadvantages of dry pipe.

The three systems mentioned above usually utilize water as the liquid suppressant. In any environment where computer equipment, specialized electronics and especially evidentiary grade electronic devices are present, due consideration should be given to the potential for water damage to technology and evidence during an event. Another consideration might be secondary Class C electrical fires spawned from a primary suppression event. In any environment that utilizes water dispersion for fire control, thought should be given to "waterproofing" concepts for certain fixtures, such as primary evidence storage. Utilizing a waterproof fire-rated safe inside the evidence locker as the primary storage container for evidence is a good countermeasure against the use of water based fire suppression systems. A fire-rated, waterproof lockbox storage system is another option for critical-to-survive evidentiary items.

Gaseous Suppression

Gas agent suppression systems, also known as Clean Agent or Total Flooding systems, provide a high end option for laboratory fire control. This class of suppressants functions in one of two ways. One group removes heat faster than it can be generated during combustion, thereby suppressing combustion. The second group depletes oxygen in order to deprive combustion of oxygen fuels. They offer advantages over water based systems such as being able to achieve total permeability in the environment. They offer advantages over chemical suppression systems because they tend to leave no chemical residues behind, lowering business recovery costs. A final positive characteristic is that these materials are in general non-conductive and they leave no conductive materials behind, making them ideal for areas with electronics. Gas suppression systems can include very complex delivery systems, and the

gas storage systems generally have a large footprint. Cost for implementation and maintenance will be high. Total flooding systems tend to require sealed environments for best effect, so other facilities costs also increase when this class of system is utilized. While these suppressants can be used in occupied space, facilities utilizing gaseous suppression should have rapid evacuation capability.

Two main classes of gas agents exist: inert gases and fluorine compound gases.

Inert Gas Suppressors

These include a number of carbon dioxide, argon and nitrogen blend gases. Inert gas suppressors are generally oxygen reducers. They tend to displace oxygen and prevent combustion via fuel deprivation. Pure CO2 suppression should never be used for laboratory fire suppression (CO2 suppression makes air completely deoxygenate and it is an active death-risk to people.) Branded suppressants such as Inergen and Pro-Inert are argon/nitrogen blends that are sold in conjunction with proprietary delivery system deployments. They can be used in populated environments. They decompose into naturally occurring atmospheric gases and they are environmentally friendly.

Flourine Compound Suppressors

These are widely utilized and they tend to be used as Halon replacements when Halon systems are upgraded. Flourine gas suppressors leach heat at a very high rate, acting as a combustion inhibitor. Branded suppressants such as Novec, FM-200 and FE-227 are common examples of suppressors in this class. They can be used in populated environments. They are environmentally friendly.

Chemical Suppression

Moving away from water dispersion and Clean Agent systems, several options for chemical suppression exist. Most chemical suppression methods require significant facilities investment to implement and significantly increase costs in many other areas of build out. For instance, hermetically sealed environments may be required when certain area chemical suppression systems are utilized. Both foam and dry chemical suppression systems are available, but both classes tend to be "messy" and inappropriate for a populated environment; such systems are generally not implemented in a data center style facility.

Electrical and Power Plant Considerations

Any high tech facility is going to have an above average power demand in order to run, cool, and keep stable all of its various technologies. In general, the cost of power provision to a forensics facility will be higher per square foot than in a "regular" corporate environment. In terms of the largest laboratory implementations, stand alone power generation facilities

and stand-by fuel tank resources may be part of the power provision plan; dedicated water provision may also be feasibly within scope for power, HVAC and even site security. In the laboratory build-out, three main categories of need should be assessed, and those categories should all be interpreted in light of both day one and growth curve demands: regular facilities load, LAN/WAN specific load, and local examiner workspace load.

The first category is the facilities load considered during every facilities build-out, i.e. the electrical demand of all general infrastructure level technology including lighting, emergency lighting, HVAC, security systems, automatic doors/windows, audio/visual implementations, telephony and communications systems, corporate equipment, general electrical consumption per employee, etc. Power provision should be generous and be cognizant of future growth as the built facility reaches 100% utilization and eventually physically expands.

The second category is the LAN/WAN load, which in any data center/forensics laboratory setting should be given independent consideration from a power perspective. Approaching the network plant according to data center grade power provision and management standards is a good base thought process. Server rooms are generally given special consideration in any build-out, but electrical provision to any network technology needs to recognize the forensics laboratory will have two fully disparate LAN provisions (business operations LAN and Examination environment LAN) and that the Examination environment LAN will need to be isolated from the general environment in terms of power provision, UPS/generator contingency planning, etc. The Examination environment LAN may also need a more robust failure/DR and redundancy plan with regard to power provision, so that it is the first environment to recover from outage and the last environment to degrade. The Examination LAN environment should, at minimum, be equipped with enough primary and secondary power for a structured, intentional safe shutdown, even under the worst external conditions. The components of power provision to the Examination LAN (and possibly all power provision) may even require special security and anti-compromise considerations, depending on the security level at which the forensics laboratory may operate.

The third category is the examination "local workspace" load. This category applies to the examination space in general and the individual examiner's functional workspaces specifically, giving special consideration to the unusually high power consumption demands per capita the forensics technical team will incur. The average corporate user group may function on a shared 20 amp circuit, powering a single workstation/monitor or laptop and a few small load items per person. A forensics investigator may well be able to max out a 30 amp circuit powering one investigation's worth of equipment, and that investigator may have numerous technology processes running concurrently in different workspaces. The examination environment of a midsize laboratory facility is likely to be "always-on" in terms of power consumption, so both environmental and equipment power consumption in the examination space will draw three times the demand experienced in the administrative portions of the facility.

Examination space needs must be assessed in terms of more than raw power consumption as well. The density and number of electrical sockets may need to be much higher in the examination space to account for the number of devices that may be active per square foot or per examination. For example, the task of cloning one hard drive may require the following devices: one forensics workstation (socket #1), the workstation monitor (socket #2), one write blocker (socket #3), one external USB hard disk (socket #4), and the original external evidence hard disk (socket #5). An investigator may have multiple cloning processes ongoing in parallel (double, triple the number of needed sockets). The ergonomics of accessing those sockets also needs consideration, favoring ease of accessibility from work surfaces. When this many devices are involved, it is important to consider not only the physical frequency of socket placement but also the density of circuit provision as well. It is important to prevent evidence grade materials from experiencing under-voltage or overvoltage conditions. Significant technical or machine time investments can be lost to a sudden power outage. Consider using a higher rated circuit in the evidence space than would be implemented in a standard corporate environment. Consider dedicated circuits per single work area. Line quality may need to be conditioned to guarantee the best integrity of the evidence hardware items. Electrical conduits in the walls may need to be shielded in order to prevent electromagnetic fields from compromising magnetically stored data in the evidence handling lanes. Transformer placement and other major electrical units need to be carefully placed on the facilities plan, shielded as necessary to abate adverse electrical fields, etc.

LAN/WAN Planning

Modeling the core technology implementation of a forensics environment on data center design is a good starting point regarding the basic requirements for a forensic laboratory technology build-out. Additional consideration needs to be given to the global and personal workspace elements of technology provision explicit to the demands of a data forensics operation.

Mention has already been made of the need to segregate the Examination environment network components from the general corporate network; in addition to the functional separation of services a number of absolute physical boundaries should also be considered. If corporate and Examination hardware is to reside in the same server room, consider a locking cage around the Examination architecture or build internal divider walls and place the Examination architecture behind a secure door: severely limit human access levels to physical space. Apply all the same security restrictions and chain of custody protocols to the Examination architecture as are applied to the Evidence room. Consider placing the Examination servers and data storage inside the Examination laboratory space proper such that all servers, data warehouses, physical cabling, switches/routers, etc. are physically protected by the same security measures restricting laboratory accesses. Route all Examination traffic through network switches dedicated to and connected physically to only Examination servers and workstations. Don't rely on virtual segregations, deploy physical segregations.

When planning the data storage needs for the laboratory facility, put emphasis on disaster recovery, redundancy, and sustainability concepts. Keep in mind that the facility needs to support large data volumes. A typical small laboratory can encounter Terabytes of data on a routine basis. Implementation of data storage for even a moderately sized facility may require an online Examination environment data storage capacity of tens or hundreds of Terabytes; this architecture will consume a significant footprint in a server room. It will be tied to other high footprint items such as large tape backup jukeboxes, nearline storage solutions, etc. Systems will need to be put in place that can handle the overhead required to maintain and preserve these enormous data volumes.

HVAC

Large numbers of computers result in enormous BTU generation (British Thermal Units, a standard measure of heat generation.) Perform very conservative calculations when determining how many tons of AC cooling are required for the technology spaces in which large amounts of heat-generating equipment reside. Make certain that cooling calculations are made from the actual equipment purchasing plans and individual device specifications, not hypothetical estimates. Keep in mind that human bodies also generate BTUs. Consider overcooling maximum capacity by a factor of 2–3x across the total HVAC design. Plan for hardware growth and factor future hardware purchases when implementing day one cooling services. Consider fully redundant units in areas that cool the Examination environment technology, and make sure either/or can provide for the entire cooling burden for the space in question. Make certain that ventilation requirements are sufficient for the spaces being cooled, and that active and passive returns are located in effective placements. If an advanced fire suppression system is in place that utilizes gas suppression, for instance, provide an active exhaust system to recover the environment once a fire event has been suppressed. Consider the water and coolant provision to any HVAC units that serve various areas; is the pipe work and pump system redundant, and does a failover system exist that guarantees the AC units will continue to be fed water? Are these feed lines protected from compromise? Are the HVAC units serving the examination space to be located over the examination space, or housed elsewhere? Placing HVAC units above the lab space adds security against physical compromise, but also adds adverse risk in the form of potential leakage and water line breakage. Environmental HVAC concerns should include noise abatement measures: an AC unit placed above the examination space may provide positive white noise in certain laboratory designs and unwelcome noise pollution in others.

Abatements

In any environment where mission critical computing systems and magnetic/tape/optical data storage reside, a number of abatement strategies need to be considered. In the forensics laboratory, most, if not all, of the following should be reviewed during the planning phase and then monitored after build-out is compete.

Temperature

All equipment has a desired temperature operating range. A typical data center will maintain an ambient temperature of 68–70 degrees Fahrenheit. Make sure the overall HVAC system can provide temperature stability within the desired ranges, even during possible HVAC equipment failures. Consider a portable cooling device standby plan. Make certain temperatures are not held at a low point that would encourage electrostatic buildup and discharge in dry air.

Humidity

Install a humidity management system that has the ability to control humidity measure to within +/− 1%. For dedicated data spaces, consider a humidity rating of 50% +/− 5%, and for occupied workspace no less than 35%. Humidity is an important factor in abating electrostatic buildup and discharge.

Static Electricity

As mentioned prior, temperature and humidity are two major environmental factors to regulate in order to avoid static electricity concerns. Give consideration to workspace elements such as anti-static flooring and actively dissipative counter surfaces and drawer linings; ground all metal furniture to earth. An operation of any size should make liberal use of portable anti-static mats and gloves, etc. Provide anti-static spray to employees wearing charge-generating fabrics.

EMI (electromagnetic interference)

Plan the electrical plant carefully to minimize electromagnetic field generation in any data storage/handling areas. Shield main power plant components such as transformers as required. Consider EMI shielding in and around the Examination laboratory space. Give strong consideration to shielding the evidence locker at a minimum. Maintain a gauss meter or series of gauss meters in the functional laboratory space, and check them regularly for anomalies. EMI regulation should speak directly to ISO planning and competency levels for any operation that specializes in electronic data handling.

Acoustic Balancing

Ambience abatements are also important in laboratory planning. Many workspaces intentionally pipe white noise into their environments in order to create acoustic masking for privacy reasons and to prevent an environment from being "too quiet"; a forensics laboratory is very likely to have many acoustically reflective surfaces, necessitating some surface texture applications, baffling, or other acoustically absorptive abatements.

Security

Security is of paramount concern to any forensics operation. Campus level access, environment level access, and object level access protocols must all be implemented. Video surveillance and live surveillance by internal security are strongly recommended. With regard to general security, the entire facility should have at a minimum a two challenge system in place such that every entrant will be providing at least one validator at an automated checkpoint (i.e. biometric entry, external security card swipe, etc.) and one other independent manual or automatic validator (sign in at security desk, internal security card swipes, etc.)

Higher levels of access control should be applied to any infrastructure or workspace related to the Examination environment or to any other environment in which evidentiary grade materials may be stored or examined. Each access attempt to the examination environment should be challenged by dual authentication and the access points should be under constant independent monitoring (i.e. cameras and access logging).

Dual authentication refers to two factor identification methodology. Two factor identification presumes any two personal identification factors will be challenged and both challenges must be successfully responded.

Challenge factors fall into the following identification categories.

- Something You Are – Biometric keys such as fingerprint or retinal scanner
- Something You Know – Password, PIN, etc.
- Something You Have – Security card, digital token, bingo card, etc.

Dual authentication across two categories of factors is recommended.

A physical sign-in/out log is a useful supplemental tool for physical plant security even if a dual authentication protocol is in place; providing an ink-signature audit trail is useful for independent audit of security system performance and original handwriting can be used to investigate identity during security audit and review phases.

Evidence Locker Security

A good locking fire-rated safe in a locked room coupled with accurate hand written access logs may prove sufficient security for a small (solo-practitioner, for instance) environment. Other evidence storage environments implement a shelf-and-cage methodology with a single portal of entry that is key locked and monitored for access. Depending on the needs of the facility and other factors, such as level of national security, the build-out of an evidence locker can become an expensive and complex endeavor.

The main security criteria to fulfill are the following:

- Is access truly restricted to the Custodian(s) of Evidence?
- Is all access to the evidence locker documented completely and without exception?

- Is all item level access (i.e. chain-of-custody) maintained correctly and without exception?
- Does an independent method of audit exist to confirm the above?

Considering security design at the corporate departmental and dedicated facility level, the highest and most restrictive levels of access control should be applied to the Evidence storage environment. Dual challenge is mandatory. Access to the evidence storage locker must be extremely limited. Only those persons with personal responsibility for evidence integrity should be allowed access. In many environments a single Custodian of Evidence is assigned master access and only that person can execute chain of custody check-ins and check-outs from the locker itself. The evidence storage environment should have dedicated security protocols for access to that environment and all accesses should be logged with 100% accuracy. Chain of custody procedures on any item entering or exiting this space should be upheld without fail. Video surveillance of the evidence storage environment is recommended with cameras on both the entry view and exit view of the door as well as coverage of the storage systems where evidence items are physically stored. An alarm should be in place to expose incursion attempts. The alarm should be robust enough to expose catastrophic entry through ceiling, walls, floor, etc. as well as unauthorized entry through the main door. The evidence storage environment should have security features built into the infrastructure itself. The walls, floor and ceiling should be hardened to discourage entry via tunneling or destruction by force; the core construction should have features such as floor-to-ceiling walls (no plenum or raised flooring, therefore no "crawl-over" or "crawl-under" unauthorized access.) Fixtures such as fire suppression and air provision should be independently controlled such that adverse events elsewhere in the facility do not cause unwanted effects inside the evidence locker itself. Air ducts need to be of a size too small for human egress and weld-grated to prevent objects from passing. No openings should be left in floor, walls or ceiling that could allow unwanted items to be inserted into or evidence items to be removed from the space.

General Ambience

As in any other professional space, the general ambience of a data forensics laboratory should be free of major distractions, providing employees an opportunity to work without disruption. The laboratory space should be a low foot traffic environment. It should be physically separated from other environments. The examination space should be well lit. The environment should promote personal comfort and positively support both standing tasks in common areas and seated tasks in personal space.

Spatial Ergonomics

A data forensics laboratory will in some ways function like a warehouse operation. The computer hard disks the forensics examiners peruse will often be provided with the rest of

the computer in tow; these chassis, monitors and other associated items will require handling and storage. Monitors, workstations, servers and other technology packages are often bulky, relatively heavy pieces of equipment. Moving bigger items to and from evidence lockdown, lifting and bending concerns surrounding transporting such items to workspaces and temporary rack system holding areas etc. should be considered during workspace design. Safety equipment such as lumbar harnesses should be made available to employees expected to execute the physical labor of lifting/carrying tasks. The traffic areas of the lab should be economized to maximize the safe execution of such tasks. Any work surface and staging area provisions should accommodate heavy vertical lifting concerns.

A Note on "common office technology"

Any evidence handling facility needs to pay special attention to potential data repositories in order to guarantee privileged information stays confidential. A forensics laboratory should include scope of such consideration to include common office technologies such as copiers and fax machines. Modern copiers and fax machines commonly have the ability to store data in memory for long periods, and either technology may have a hard disk on board! Maintenance plans for such devices should consider the possibility of privileged information being resident, and security and audit methodology should be applied to guarantee proper handling/destruction of any storage medium's contents.

Personal Workspace Design

Each laboratory inhabitant should be provided a significant amount of operating space. Work surface area should be bountiful, especially digital work surface area (i.e. monitor footprint). Electricity supply should be robust. The personal space of each examiner should be considered a "mini-laboratory" and that mini-lab should be stocked with all the hardware and software necessary for an examiner to perform common investigative tasks and to maintain work product. A dedicated investigation platform, a complete kit of write blockers and accessories, a separate system for corporate/business communications, a workspace level data management system and a close-at-hand library of reference materials are all desired elements for an active and utile personal investigation workspace.

Common Area Considerations

Consider providing multiple units of every technology. Multiple sets of write blockers and multiple investigation machines allow for several parallel forensics tasks to occur. Design workspaces with a template design to allow multiple individuals to execute similar tasks concurrently in different workspaces, or to allow one individual to rotate between several stations in order to manage multiple machine-time intensive tasks. Design work areas to support the execution of multiple tasks with minimal foot traffic. Deploy sufficient "shared resources" to effectively serve the needs of staff without causing workflow bottlenecks

(for instance, when scoping DVD production capability make sure the DVD burning tower has a job scheduling capability in order to capitalize on a full 24 hour production cycle regardless of staffed shift availability. Determine how many DVD burning towers are required to fully serve departmental needs. Is one large central unit the best choice, or are four smaller units located in different areas the most effective option? Etc.)

Essential Laboratory Tools

The tools of the trade: essential and specialized technology for both Field and Laboratory (see Table 1.1).

NOTE

The author is not attempting to endorse the use of any specific product, or to be exhaustive in the description of capability or utilization of any product listed herein. This chapter hopes simply to expose the reader to a wide selection of readily available hardware and software tools.

Table 1.1 Vendor Reference Matrix

Vendor	Product Examples	Web Site
Tableau	Write block devices	www.tableau.com
Intelligent Computer Solutions	Hardware, Write block devices	www.ics-iq.com
Wiebetech	Write block devices, Hardware	www.wiebetech.com
Digital Intelligence	Write block devices, hardware, software	www.digitalintelligence.com
MyKey Technology	Write block devices	www.mykeytech.com
Guidance Software	Write block devices, software	www.guidancesoftware.com
Paraben Corporation	Write block devices, hardware, software	www.paraben-forensics.com
Forensic Computers com	Hardware, Software, Systems	www.forensic-computers.
LogiCube Forensics	Hardware, Software	www.logicubeforensics.com
VOOM Technologies	Hardware	www.voomtech.com

Table 1.1 Continued

Vendor	Product Examples	Web Site
DIBS USA	Hardware, Software, Systems	www.dibsusa.com
Fernico	Software/Hardware	www.fernico.com
Primera Technology	Software/Hardware	www.primera.com
Project-A-Phone	Hardware	www.projectaphone.com
Rimage	Hardware	www.rimage.com
Ashby	Hardware	www.ashbyind.com
CopyPro	Hardware	www.copypro.com

Write Blockers

No laboratory or field forensics toolkit would be complete without write block methodology and devices. Prevention of data spoliation (the compromise of data integrity by intentionally or inadvertently altering the state of the data from its "original" form) is a prime directive for forensic examiners. The courts will challenge forensics work product in order to leverage spoliation concerns; one of the most common attacks on forensics work product focuses upon the methodologies employed when handling digital evidence. Was the evidence maintained in an original state? Were the conclusions drawn based upon uncompromised materials? Were the proper tools used in the process? Etc. In a vast number of circumstances, when an unprotected writeable data device is connected to a computer, it will incur change. Computer boot sequences, volume mounts, and a plethora of other events can modify some component of the evidence data store if it is not explicitly protected from write-to events. A forensic examination environment, therefore, will host a broad range of methodologies and devices insuring write block capability.

In certain circumstances, utilizing the proper methodology, it is possible to achieve a no-write status at a software level. Certain windows registry edits can protect USB devices from write events; Linux volumes can be mounted in a certain way as to make data stores read-only. Microsoft DOS and Linux operating systems can both be modified to be forensically sound with regard to data stores, and they can be packaged as a self-contained bootable environment from diskette, CDR, DVD, thumb drive, etc. Of course, an investigator implementing these techniques must repeatedly test their methodology and be able to both demonstrate and explain proofs that the methods truly are forensically sound in order to defeat the inevitable court challenges to method that will arise.

Hardware write block devices (aka write blockers, forensic bridges) are a flexible, extremely utile, and core component of the forensics tool kit. They have the advantages of portability, broad application, ease of use, and ease of function testing, etc. It is a simpler task to visually and conceptually demonstrate the function of a hardware write block device than

to explain the esoteric technical specifics of a Linux operating system forensic recompile to a jury of average (and often non-technical) citizens.

A number of major write block devices are readily available for purchase. Common write block brands include Tableau, WiebeTech and Intelligent Computer Solutions DriveLock. Digital Intelligence carries the complete line of Tableau products under the brands UltraBlock and FireFly. Guidance Software has FastBloc, based upon WiebeTech firmware. Paraben Corporation vends LockDown. MyKey Technology vends NoWrite. A majority of these write blockers are designed to be portable, allowing them to have equal value on a laboratory bench or in the field.

Multiple form factors exist to serve different environmental needs. Hard disk technology has multiple interface types (IDE, SATA, SCSI etc.) Write block technology therefore integrates multiple interface types in order to attend to the diversity of connectivity needs an investigator may encounter. Many write blockers are designed to support the IDE hard disk interface and will have an adaptor/cable kit that allows the device to also support SATA. Tableau has an explicit forensic bridge model for each interface type: SATA (T3u and T15), IDE (T5 and T14), SCSI (T4 and adaptor kit), USB (T8), etc. There are advantages to both design scenarios. USB and FireWire form factors are common interface types utilized to connect external write blockers to examination machines. Most write blockers will be packaged with the appropriate power supplies and cabling to support that device.

Forensic bridges can also be acquired for permanent installation into workstations. While not portable (unless the entire workstation is portable), the internal forensic bridges usually have the advantage of being space efficient, often providing several evidence drive interfaces while only consuming one device bay on the workstation. The Tableau T35i Combination Bridge is an example of a permanent mount device.

Write block technology also exists to support examination of non-hard disk media. A number of multi-format forensic card readers exist to handle SD, SDC, xD, MMC, CF, etc. Examples include the Tableau TDA8-M 12-in-one reader, the Addonics DigiDrive 12-in-1 Flash Media Reader, and the UltraBlock Forensic Card Reader from Digital Intelligence. In addition to the core write block device, passive format-to-format adaptors can be purchased from retail and specialty outlets that can adapt a non-supported card format to a supported card format, further enhancing the interface capability of the multi-card readers. (note: when adapting formats always test the adaptor to insure it is indeed a passive, non-change inducing device).

Write Block Field Kits

Forensic Bridge field kits are an excellent addition to the forensic laboratory inventory. Aside from the obvious (field use), a field kit can be fully functional on an examiner's laboratory desktop and as such help reduce inventory purchase costs by minimizing the amount of hardware per examiner required to execute data acquisition and investigation in diverse environments. Field kits tend to be lightweight, ruggedized, designed to meet

air transport criteria, and packed with device, adaptor and cabling options in order to address as many of the "unknowns" of field work as possible. The Digital Intelligence UltraKit and the Ultimate Forensic Write Protection Kits from Forensic Computers are excellent examples of single-package systems; they include a majority of the Tableau devices referenced prior as well as numerous supporting parts. Field kits also commonly supply a basic multi-function hand tool kit, bit/driver set and a digital camera in order to help support other aspects of field work. A good core field kit can be substantially fleshed out with cabling, adaptors, extra devices, etc. to create a very powerful and economical portable laboratory system. Always include redundancy for high use/fragile components: multiple AC adaptors, power cords, and interface cabling units are a must. Convenience items such as the Tableau in-line power switch (T2) add a level of protection to the examination equipment assembly process and help protect against damage to evidence media via pilot error.

One major implementation of write block methodology focuses upon protecting original media from change during examination. Another implementation is protecting original media from change during duplication. In many instances, field investigation practices will require the acquisition of data from the wild for later study. This acquisition often occurs via the creation of a forensic duplicate of original evidentiary materials for transport of evidence back to a laboratory environment for analysis. In such cases, the need for write blocking will be conjoined to the need for a duplication platform.

Hardware Duplication Platforms

A number of handheld and desktop forensic duplication systems are available. The core functions that they provide include write blocking the original evidence media, performing a data duplication process to secondary media, and measuring the correctness/completeness of the duplication process via some measurement criteria (almost always via utilizing a hash algorithm such as MD5 or SHA1 or both) to validate that the entire original was duplicated to the forensic copy. Several devices of this class also integrate reporting capability.

A number of popular models are readily available. The Logicube Forensic Talon, boasting a data duplication rate of up to 4 GB per minute, provides multiple media adaptor kits and possesses extensive reporting capability. Intelligent Computer Solutions ImageMASSter Solo-III forensic duplication device handles several interface types and has the ability to write to two output hard drives concurrently. Voom Technologies Hardcopy II provides a simple interface and handles IDE hard disk duplication (expandable to SATA duplication with adaptors). Voom Technologies also produces a SCSI HardCopy for SCSI platform acquisitions. Some vendors package multiple hardware duplication devices and accessories into a field kit: the DIBS RAID: Rapid Action Imaging Device is an example. Many of these devices also provide the output options of bit-for-bit duplication, one or more forensic

image format acquisitions, and transport media sterilization. Hardware based duplication platforms tend to have much faster data transcription rates than software based duplication solutions.

Duplication hardware is a valuable addition to the examiner's toolbox, but duplication tools generally provide no environment in which an examiner can investigate the data being duplicated. Portable forensics computer systems provide an examiner with an investigation environment, expanding the examiner's field capability one step further.

Portable Forensics Systems

When the need to take the entire investigation process into the field arises, a forensic examiner must have access to not only the protective and duplication tools available, but also fully interactive examination environments. Specialized portable forensics computing systems provide a highly mobile, equipment intensive and methodology sound platform for the forensics examiner. Complete field examination systems allow the examiner the ability to duplicate digital evidence and analyze said evidence on one robust platform. "By-hand transport" level portable forensics systems will generally be provided in the form of ruggedized, feature packed laptops or custom "suitcase style" workstations. A second tier of "portable" forensic systems includes a class of machines and mini-networks that are heavily ruggedized for mobility but are not intended for day-to-day high mobility. All of these investigation systems tend to have fast processor capability, copious amounts of memory and high volume data storage space. Most will be optimized for running specific forensic software packages. Implementation of multiple operating systems on one workstation is common.

Every examiner will need a personal field kit inventory that is easily manageable for loading/unloading into vehicles and for air travel. High mobility portable systems are often designed to rely on external field kits such as the write blocker field kit and supplemental cable and adaptor solutions in order to make the core system smaller and more transportable. The below matrix references a number of such systems and examples of core features:

NOTE

The author does not intend for the feature lists shown in Table 1.2 to be exhaustive representations of any given system's hardware or software specifications. For comprehensive detail refer to the web site or product documentation explicit to the system in question.

Table 1.2 Laptop style portable forensics solutions

Vendor	System	Class	Feature Examples
Forensic Computers	Forensic Air-Lite VI MKIII	Forensic Laptop plus external bridge kit	2 GB memory; Supplied with write blocker accessory kit; 25-in-1 media card reader; ruggedized hardshell case; extra external HD storage supplied
LogiCube	"PFL" Portable Forensics Lab	Forensic Laptop plus external duplicator kit	2 GB memory; Supplied with a Forensic Talon drive duplicator; bundled with forensic investigation software suite FTK from Access Data; ruggedized hardshell case
DIBS USA	DIBS Mobile Forensic Workstation	Forensic Laptop Plus external accessories	1 GB memory; Supplied with an Inkjet printer in hardshell case; bundles with pre-installed forensic software including Access Data FTK, CD/DVD analysis tools and other forensics utilities; digital camera; write-block accessories
Digital Intelligence	FRED-L	Forensic Laptop plus external bridge kit	2 GB memory; Multiple bootable operating systems onboard; Supplied with UltraKit write block kit; forensic card reader; ruggedized hardshell case

"Workstation-in-a-box" style computers can offer a few flexibilities to field personnel that may not be available in the laptop style kits, such as a higher number and/or more friendly form factor of available slots for add-on devices. The suitcase style workstations often have a detachable monitor/keyboard/mouse set that can be used to work with evidence workstations (presuming those components are not available at the field site) for boot-up procedures such as BIOS checks and verifying proper suspect system reassembly (see Table 1.3).

Table 1.3 Suitcase workstation style portable forensics solutions:

Vendor	System	Class	Feature Examples
Forensic Computers	Forensic Air-Lite IV MKII	Forensic Workstation (suitcase type)	Pentium IV 4.3 Ghz; 2 GB memory; external keyboard & mouse; ruggedized hard-shell case; LCD monitor; upgradeable
Digital Intelligence	FREDDIE	Forensic Workstation (custom design)	4 GB memory; Multiple bootable operating systems onboard; two removable hard disk bays; multiple onboard write block formats; supplemental toolbox including hand tools and camera; optional hardshell case; multiple accessories and several onboard software utilities

Portable Enterprise Systems

In some instances field portability concerns address the need for a robust, temporary laboratory facility at an examination location. Forensics portability can be extended to "network-in-a-box" solutions. "Half-rack" solutions can fill this need. A portable enterprise system will offer core components such as the examination system(s) plus integrated write block bridges plus robust examination hard disk storage space (multi-Terabyte RAID level storage, etc.) plus all add-on hardware such as monitors, kvm, etc. wrapped into one ruggedized cage. This type of portable environment is usually high durability but low mobility, quite weighty and meant to be transported crated and packed (i.e. set-up and break-down time) as opposed to the more "plug-and-play" high mobility equipment.

Laboratory Forensics Systems

In many instances, the high mobility equipment that accompanies the field examiner on excursions can also be utilized on the desktop. Small operations, especially solo practictioners, will find great economy in purchasing their primary gear with both field and desktop implementation in mind. For facilities that can support a permanent lab installation of desktop investigative gear plus field support equipment, numerous "non-portable" investigative powerhouse systems are available. These systems tend to offer all the various field hardware

solutions found in portable kits for write blocking and hard disk management combined into one desktop chassis. A solo practitioner or cost-conscious operation may find value in a "white box" approach, building a desktop system from scratch to suit specific needs; all-in-one devices such as the Tableau T35i Combination Bridge and Tableau T335 Drive Bay Controller are economical options for implementing multiple write-block and multiple hard disk solutions in a single chassis.

Pre-built desktop forensics systems will often have the best computing power available at time of purchase (portable technology historically tends to lag behind desktop technology in terms of "bigger-better-faster-more").

When you are choosing the specifications for a desktop laboratory processing system, always target the fastest processing, largest memory allocation and largest possible hard disk drive volume available at time of purchase, subject to any budget constraints. Hardware ages quickly, and maximizing the point of purchase processing capability of new systems maximizes the useful lifespan of the equipment. Given the process intensive needs of most forensic software application suites, a fast, powerful CPU (central processing unit) and a large amount of RAM (random access memory) are critical.

Typical data storage space requirements for the forensics examiner are astronomical. As of the writing of this chapter a single hard disk of one Terabyte volume is readily available in the retail consumer marketplace and a 1TB hard disk is commonly available in home and corporate computers at point of sale. Encountering hard disks of 500 GB or 750 GB is commonplace. Current pre-built forensics systems offer BASE storage of 2–6 Terabytes per system. Make maximizing storage space a priority and consider the relatively short span of any volume's sufficiency when allocating resources to acquire forensics computing equipment. Evaluate hardware level redundancies and robust backup systems for managing data volumes of this volume.

Provide considerable monitor real estate. Forensic examiners have to visualize enormous amounts of data during an examination, and the viewable area of the computer monitors upon which they work can have notable impact on investigation speed and efficiency. Many forensics systems are sold with dual-head video cards, such that two or four monitors may be attached to one system. 22"-30" flat panel monitors are space efficient, readily available and reasonably priced.

The ability to operate under multiple operating systems is greatly desired. The forensics examiner will make use of both cutting edge and "old school" investigation tools, requiring multiple operating systems to support those tools. From an investigative standpoint, the investigator will routinely analyze evidence derived from (and will therefore require access to) multiple operating system environments. Provision of multiple operating systems extends across version levels (such as multi-boot options for Microsoft DOS 6.22, Windows 98 SE, Windows 2000 and Windows XP) and platforms (Windows and Linux boot options). The more powerful pre-built forensics systems will provide four or more bootable operating systems.

Media Sterilization Systems

Spoliation challenges to evidence integrity include challenges to the duplicate evidence copy; often an argument will be made charging evidence could experience spoliation due to data artifacts pre-existing on hard drives used as duplication destination. In order to help prevent questions of this type from arising, a solid policy for work product media sterilization should be in place for any forensics practice. Any hard drive to be utilized as substrate for an evidence duplicate should be sterilized prior to use and documented as sterile. Furthermore, the sterile state of such media should be validated by some post-sterilization procedure. Some forensic hardware and software duplication tools sterilize in conjunction with data acquisition by hash validating written sectors and subsequently zeroing out all other writable space; they will validate the acquired evidence data stream via hash methodology, then "wipe" any remaining writeable data space to a random or zero value via data overwrite methods. Relying on this process is more complicated from an "explain-to-jury" standpoint than confirming all new substrate media as clean and viable for use, prior to use.

Software solutions such as Guidance Software's EnCase Forensic examination suite include the capability to sterilize and subsequently validate hard disk media; retail products such as White Canyon Software's WipeDrive (www.whitecanyon.com) can destroy data according to a number of data overwrite patterns. (note: MANY retail "drive wiping" programs are ineffective at complete data destruction – if any software data destruction method is utilized make certain to test results and case-by-case validate any destruction attempts). Hardware sterilization devices exist that can bulk-overwrite hard disk media.

If any evidentiary or sensitive data needs destruction after its value expires, a sterilization process and validation process should be applied to the media in order to destroy data. Both hardware and software sterilization tools exist that can destroy data. A forensic laboratory environment will often need to demonstrate adherence to commonly accepted practice. Data destruction practices are no exception. One of the most popular courtroom attacks on digital evidence centers around spoliation; it behooves a forensics practice to not only practice good method but also adopt commonly defined "industry standard" practices with regard to work product media consumption. While an argument can be made that no truly "industry standard" definition for data destruction exists, several published standards do serve as common reference materials for most forensics practitioners on the topic of data destruction practices.

By far the most referenced document is DoD (Department of Defense) 5220.22-M. Commercial practices especially, in order to lend market credence to their vended product, will claim adherence to "DoD grade" destruction practices. This document, reissued in February of 2006 as the NISPOM ("National Industrial Security Program Operating Manual") is, in its current incarnation, a high level document and it does not speak to actual technical specifics of sanitization method. Prior to 2007 the Department of Defense CSM ("Clearing and Sanitization Matrix") directly defined the term "sanitize," and it included

specific technical processes for n-pass overwrite techniques on magnetic media as viable sterilization techniques (including a three pass-plus-verification data overwrite technique commonly cited as the DoD 5220.22-M standard). In June of 2007 an updated CSM was published by the Defense Security Service of the DoD explicitly disavowing the process of data overwriting as a valid method of sanitization for classified materials; the new matrix endorses degaussing and physical total destruction techniques (incineration/chemical abrasion) of magnetic media to truly "make data gone". (http://www.dss.mil/isp/odaa/documents/clearing_and_sanitization_matrix.pdf)

Other United States government derived "industry standards" include the 1993 NAVSO (Navy Security Operations) P-5239-26 MFM/RLL standard for dealing with data remanence on hard disks using MFM/RLL encoding schema. This standard continues to endorse degaussing and data overwrites, but it is currently best interpreted in light of DoD 2007 CSM.

The UK governing body standard "HMG Infosec Standard No 5" is another storage device disposition standards document gaining popular citation; various other data overwrite standards and protocols are frequently cited, including GOST P50739–95 (Soviet) and Gutmann-35 (New Zealand). Data destruction services may also include references to procedural HIPAA (Health Insurance Portability and Accountability Act) compliance regarding data storage media that also happens to hold content falling under the scope and provisions of the Act.

Ultimately, a forensics laboratory should have data destruction tools on hand to address the real needs of that facility; a solo practitioner may never need access to an incinerator or hardware shredder. Environments engaging in data destruction practices as a vended or self service should consider a bulk data wiping hardware device for new media preparation and a degauss chamber or physical destruction device for any disposal sterilization needs.

Data Management (Backup, Retention, Preservation)

Forensics operations have a significant need for data storage and preservation. When considering implementation of a forensics environment, consider the many touch points for data storage and data transactions that typically exist. Forensic examiners require high volumes of digital workspace; examination workstations may have Terabytes of onboard storage, and high volumes of data may transact across that environment on a regular basis. This workspace needs to be managed as part of the examiner's workflow. Common network data stores will also be in the Terabyte (or higher) storage measures. Networked storage may have data content requiring long term preservation, rapid workflow turnaround to archive, or extended presence in on-line storage. A robust policy of data handling strategies will need to be developed to accommodate these different levels of need. A forensics facility, given the evidentiary nature of original and work product data and hardware, will require far more

stringent policy and procedure for oversight, audit and reporting on data management activities as compared to the typical corporate or data center operation. Data management solutions provided at both "per machine" and enterprise level guarantee flexible options for efficient data handling, potentially minimizing workflow based bottlenecks.

Modern hard disk sizes indicate the need for high volume data storage solutions. Even a solo practitioner can reasonably expect to have a need to manage Terabytes of information. A forensics environment may want to consider a number of options to ameliorate the need for constant, rapid, and high volume data management. Possible components for data management include various hardware and software solutions for data storage and movement. Tape backup systems deployed at a "per machine" and enterprise level provide the ability to back up and store large volumes of work product via "clone-validate-delete" methodology. Near-line storage in the form of redundant or low-speed disk storage or tape jukeboxes can help serve the offloading and retrieval needs of high volume data management by providing "stand-by" access levels to low or intermittent use work product. Offline storage (tape, optical disk and offline hard disk arrays) allow multiple options for long term preservation. Hot-swappable/external storage devices add value to the overall data management proposition by providing "instant access" or "instant storage" methods for data repository refreshes and fast data offlining.

Various magnetic tape and DVD ("Digital Versatile Disc") formats are readily available and cost effective form factors for evidence preservation. Magnetic storage such as hard disk or tape can provide high volume storage solutions when environmental factors are tightly controlled; both media formats are magnetic and subject to limitations regarding data degradation over time.

Magnetic tape offers the benefit of tamper resistance: to add or delete data results in obvious physical signs (splicing) or electronic signs (readability/non-readability.) Downsides to tape use include backup software versioning changes adversely impacting tape readability, relatively little recourse for restoration when a tape span media tapes fail, and simple media degradation over time. Tape hardware tends to be expensive and can be a significant investment at an enterprise level. Installation and licensing for backup software at an enterprise level also tends to be a significant cost of business. Even so, tape systems, even high end hardware rotary jukeboxes and advanced library software systems, tend to be more cost effective than full hardware storage redundancies such as parallel SANs (Storage Area Networks).

CD/DVD Hardware Solutions

Many long-term evidence preservation and storage needs are addressed via the use of optical media such as DVD or CD (Compact Digital disk). Optical media tends to be very inexpensive and readily available. Use of CD media as a long term storage media was commonplace prior to hard disk space and magnetic tape hardware becoming cost effective to purchase and

utilize; CD and DVD media still have practical application as back-up or unitized evidence storage, especially with newer DVD format storage capacity improvements. High grade optical media tends to have a life span measured in decades, making it ideal for long term evidence preservation (caveat: make sure your hardware standard will survive as long as the media needs to survive!). For instance, one common evidence acquisition methodology involves the creation of "image file sets" of a specific format and size: most forensic software and hardware offerings supporting image set creation also allow the size of image files to be dictated by the examiner. Default image segment file sizes tend to reflect the anticipated use of CD as the ultimate storage unit (i.e. 650 MB per image file, allowing one image segment to be written to one CD media).

Data and media duplication hardware has been developed to support the special needs of a forensics environment. One example of such technology is the Fernico FAR system. The FAR system is based upon CDR/DVD burning technology from Primera Technology and is augmented by specialized forensic data archiving software written by Fernico. The combined system provides a robust data management and archival tool that provides several services to the forensic examiner that are explicit to evidence management needs. The system provides the benefits of automation to a typically tedious and manpower intensive series of tasks. The FAR system will burn DVDs or CDRs, label the media, intelligently manage the spanning of data across disks, and apply forensically desired validators such as robust hashing options to the data burned to disk. The device is network capable, and it is able to automate activities utilizing a calendar and job based scheduling system. The system is designed to not only archive data to disk but also restore data from archive libraries of disks. A FAR system may also be utilized to correctly duplicate CD/DVD original evidence media. Multiple FAR systems are available, each catering to a specific scope of media capacity and disk production. Other robust optical media production systems exist for high volume CDR and DVD production, provided by vendors such as Rimage. Whether designed explicitly to serve forensics needs or simply to serve as a vehicle for work product duplication and publication, an optical media production platform is a valuable addition to the forensics laboratory tool set.

Disk duplication systems are not limited to optical media. Forensic laboratories must be prepared to process older storage technologies as well; floppy diskettes are still a viable media to encounter and evidence duplication services for this type of media can be automated by bulk duplication machines as well. Vendors such as Ashby and CopyPro provide floppy diskette duplication systems that auto-feed bulk diskette media and bit-for-bit validate copies.

Portable Device Forensics, Some Basic Tools

Faraday Devices as Applied to Forensics

From both an information privacy and data integrity standpoint, the science of portable device forensics embraces the use of various configurations of the Faraday enclosure in order to preserve data integrity and/or prevent data transmission.

A Faraday enclosure is simply an enclosure capable of blocking electromagnetic fields and energy waveforms. A Faraday enclosure is composed of conductive materials (layers of variously composed wire mesh, for instance) that allow for electrical induction to occur across the surface of the materials when energy is applied to that surface. When a Faraday enclosure is introduced to a penetrating electromagnetic field or a waveform of a particular frequency or range of frequencies, the applied field/waveforms will not penetrate the enclosure surface; they will instead travel across the conductive surface(s) of the enclosure.

If a signal generator is placed inside a Faraday enclosure, that signal can be "kept inside the box."

If a signal receiver is placed inside a Faraday enclosure, the enclosure can prevent the receiver from "hearing the signal."

Real-World Examples

RFID (Radio Frequency Identification) tags that actively transmit information are becoming widespread in actual/planned implementations of personal data management. Faraday systems have been designed specifically to defeat RFID scanners picking up signals from active RFID tags inside the Faraday enclosure. Conversely, the enclosure can also prevent active RFID scanners from successfully "pinging" a passive RFID tag. Paraben Forensics vends forensic products that explicitly address RFID data management. (see www.paraben-forensics.com for an applied example of the Paraben Passport StrongHold Bag product referencing RFID technology applied to future versions United States Passports.)

Cellular telephones are ubiquitous, protocol chatty data storage repositories. Forensic acquisition methodology focuses not only upon the data they may directly harbor, but also the air gap protocol channels via which they communicate. External signals capable of compromising data integrity (spoliation via inbound signals modifying/deleting data) are blocked, and internal signals are prevented from reaching intended destinations. In order to islolate a device that may be actively in communication with cellular towers, other proximate devices, etc, the use of a Faraday enclosure is paramount. Another product in the Paraben StrongHold line, the StrongHold Bag, is designed for convenient and highly portable field isolation, transportation, and storage of cellular telephones and other transmission/receipt capable portable devices. Guidance Software offers WaveShield bags, a similar signal blocking solution (www.guidancesoftware.com).

Other applications of Faraday enclosures include both portable tent and sealed box environments that can be used for field analysis, isolation and portage of suspect devices. The Paraben Wireless StrongHold Tent and Wireless StrongHold Box are both designed to be fully portable Faraday environments enabling the isolation and examination of portable devices in the field. Even an entire Forensics laboratory environment can be shielded inside a Faraday enclosure! Specialized environmental abatement companies cater to this facet of facilities buildout, designing anti-electromagnetic shielding for many industries, including medical, military, and of course, data forensics environments.

Portable Devices and Data Storage

Locating the Data

Data on portable devices can reside in many storage forms: SIM cards and other types of card media, chip-based storage, etc. are common, and modern portable devices will typically have more than one of these storage mechanisms in place. In addition to the multiplicity of data storage form factors, the form factor may predispose either static data storage (for example, phone numbers stored on a SIM card) or volatile data storage (recently dialed phone numbers stored in the battery powered, chip based memory).

Power

First rule of securing portables for forensic examination: if the device is off, leave it off (but keep the battery charged); if the device is on, do everything feasible to keep the device powered on until a full examination can be performed. In order to help power management and "limited time to respond" portable device research, a number of tools can be added to the examiner's toolkit.

Multi-purpose power devices such as Paraben's Remote Charger and Handheld First Responder Kit (includes power device, faraday bags and portable device flash cards) supply power to a device to significantly extend its' data preservation window, giving the forensic examiner time to transport the device to a laboratory environment for detailed analysis.

For assistance in documenting a live review process in lab or field, tools such as Project-A-Phone device analysis platform and the Fernico ZRT Mobile Device Screen Capture Tool help examiners do rapid video examination and preservation of "power on" screen states, menu system function and other live-system processes that are difficult to recover by a data export process or to document for examination and later presentation purposes.

Readers, readers, readers!

Many of the card types inherent to cell phone technology (notably SIM) are proprietary – in order to decipher SIM contents, custom forensics software/hardware packages are usually required. For the more common data storage card types (common form factors for PDAs, cameras, cell phones, etc. such as SD, MMC and CF formats) a good 15-in-one card reader is a good tool to add to the toolkit; forensically sound card media readers are distributed via forensics vendors such as Digital Intelligence (see: UltraBlock Forensic Card Reader on the Digital Intelligence web site.)

Cables, cables, cables!

A good portable device forensics toolkit will include many, many portable device power and data cables. Many of the all-in-one forensics cell phone and PDA examination suites provide an extensive cable selection for both powering and linking to the multitude of data interface

formats found on portable devices. Don't discount the retail cell phone outlets, retail electronics stores and web sites for supplemental cables and adaptors to augment your mobile device field kit! In addition to cables, cables, cables, a good toolkit also includes adaptors, adaptors, adaptors!

Several expansive portable device forensic toolkits are available, including:

- The Paraben Device Seizure Toolbox www.paraben-forensics.com
- LogiCube CellDEK www.logicube.com
- Guidance Software Neutrino www.guidancesoftware.com
- Secure View Forensics www.datapilot.com

NOTE

For an up to date general principles guide to Mobile Forensics, and an excellent listing of supplemental resources, terms and tools, please refer to the NIST (National Institute of Standards and Technology) Special Publication 800-101, "Guidelines on Cell Phone Forensics", released in May of 2007. http://csrc.nist.gov/publications/nistpubs/800-101/SP800-101.pdf

Forensic Software

The realm of forensic software selection is broad. A forensic examiner will require a diverse array of applications in order to effectively address the myriad investigative scenarios encountered during course of business. It is important to understand that no one tool is all-inclusive, therefore having multiple tools on hand is desired. Many of the tools may have overlapping functions, but each application will have unique strengths that add value to an overall investigative process. Having multiple tools on hand that perform similar functions allows processes and methods to be validated. The unique strength of each tool applied to a given problem will allow for a diverse and thorough investigation to take place. When provisioning software services to an investigative team, it is important to consider a few main areas of focus.

- Operating Systems
- File Systems
- Main Investigative Platform
- Special Focus Tools
- Tools in the Enterprise
- Ad Hoc scripts and programs

Operating Systems

Operating systems are software implementations. Miriam Webster Online (www.m-w.com) defines "operating system" as follows: "software that controls the operation of a computer and directs the processing of programs (as by assigning storage space in memory and controlling input and output functions)." When considering any forensic laboratory implementation, consider the need for possessing a software library containing a large number of operating systems, including any available version levels of each operating system retained.

A forensic examiner will want to leverage multiple operating systems as his/her production environment, i.e. the platform upon which he/she conducts an investigation. The reasons for wanting a good library of operating systems in a forensic laboratory are many. From a provision of services standpoint the ability to work with multiple Microsoft Windows versions, Linux/Unix flavors, and Apple/Mac OS versions should be standard. Numerous other operating systems exist as well.

Leveraging multiple operating systems allows the examiner to use software applications written for specific operating system platforms. For instance, a twenty year old software application may provide an examiner with a needed function that no newer software provides, and in order to run that special program the use of an operating system from the same era may be required.

The decision regarding which operating system to use for a given examination may be strategic. A forensic examiner may want to examine an NTFS format hard disk from a Windows operating system suspect computer while working on an examination system running one of the Linux operating system flavors in order to defeat a spoliation argument such as "your Windows examination computer made those Windows artifacts, changing the evidence". A plethora of other proper handling best practices aside, the examiner's response to such a challenge could subsequently be "Linux doesn't create Windows artifacts".

A forensic examiner may be called upon to examine suspect computers configured with a diverse spectrum of operating systems: having a library of installable operating systems allows the investigator to create a test environment and/or reconstruct events in the same software universe as the suspect computer.

File Systems

Different operating systems support different file systems; in order to examine file systems of certain types, it may be necessary to have a compatible operating system available in order to do so.

File systems are not software per se; a file system is a method for storing and organizing data content in/on a given medium. The most common application of the term "file system" in computer forensics usually refers to the organizational structure of electronic computer data stored on computer media such as hard/floppy/optical disks, thumb drives, etc. A large number of file systems exist. A forensic examiner will want to be prepared to work upon as

many different file systems as possible: an examiner will need access to multiple operating systems, multiple hardware devices and various software applications in order to attack a broad scope of file systems.

Investigative Platforms

Multiple software applications fill the role of "main investigative software platform" for a forensics examiner. Common characteristics of the main investigative tool suite include the ability to create an "image file set" or "bit for bit clone" copy of suspect media, analyze live or imaged datastreams, search and cull suspect data content, generate reporting on contents and findings, and export data of interest for subsequent use external to the software. Investigative software suites are available for multiple operating systems (Windows OS, Linux OS and Mac OS). Some investigative software is restricted to law enforcement (iLook) while other application suites are available to the general public (EnCase Forensic, Forensic Tool Kit, S.M.A.R.T., MacForensicsLab).

Table 1.4 Vendor Reference Matrix

Vendor	Product	Web Site
Guidance Software	EnCase Forensic, EnCase Enterprise	www.guidancesoftware.com
AccessData Corporation	Forensic Tool Kit (FTK), Password Recovery Toolkit (PRTK), Registry Viewer	www.accessdata.com
IRS-CI Electronic Crimes Pgm	iLook** **Law Enforcement ONLY**	www.ilook-forensics.org
SubRosaSoft	MacForensicsLab	www.subrosasoft.com
ASR Data Acquisition & Analysis	S.M.A.R.T. for Linux	www.asrdata.com
Paraben Corporation	Chat Examiner, Email Examiner, Device Seizure	www.paraben.com
Hot Pepper Technology	EMail Detective	www.hotpepperinc.com
Wetstone	Gargoyle Investigator; LiveWire Investigator	www.wetstonetech.com
TechPathways	Prodiscover Forensic/IR	www.techpathways.com

Each of the aforementioned investigative suites has a unique set of strengths. A forensic examiner often possesses several software suites and leverages more than one investigative suite during an investigation. EnCase Forensic by Guidance Software has a robust scripting

language empowering custom investigation. Forensic Tool Kit from AccessData has a very intuitive user interface and the ability to integrate password breaking software from the same vendor. MacForensicsLab is currently the only major forensic investigative suite developed to run on an Apple OS X system; S.M.A.R.T. is a robust tool suite developed for the Linux operating system. ProDiscover Forensic by Techpathways is capable of performing examinations on Sun Solaris UFS media.

Other/Specialty Tools

Many software packages have been developed to attend to specific data analysis needs. Email analysis tools, password breaking tools, decryption tools, portable hardware analysis tools, artifact specific identification and analysis tools, and **many** other types of data identification, conversion and analysis tools have been developed, all of which may prove valuable to the forensic data examiner. AccessData's Password Recovery Toolkit executes password breaking and certain decryption tasks. Access Data Registry Viewer enables Windows registry examinations. Paraben's Chat Examiner, Email Examiner and Device Seizure all provide specific realm analysis capability. Hot Pepper Technology's EMail Detective AOL email analysis provides product specific email analysis. Wetstone's Stego Suite provides Steganographic investigation capability; Wetstone's Gargoyle Investigator provides malware detection. These products and many other special focus tools can supplement an investigator's digital toolbox. An examiner's digital toolbox should be as robust as, if not more robust than, their physical toolbox.

Tools in the Enterprise

Forensic examination may also be executed against local and wide area network resources in the enterprise. Whether responding to an acute need for examination across a corporate wide area network or functioning as a forensics service permanently attached to a corporate security group, investigation tools have been developed to address forensic grade data harvesting and reporting in widely dispersed environs. Enterprise forensic tools often possess capability to investigate live systems remotely and analyze volatile memory contents and network metrics as well as local machine activities in situ.

Guidance Software's EnCase Enterprise edition has all the functional aspects of EnCase Forensic regarding data collection, analysis and reporting: the enterprise version allows those functions to be driven across networked resources via use of local application connector clients and centralized investigative resources. Wetstone Live Wire Investigator also allows data collection, analysis and reporting over a network of distributed resources. Live Wire is an "agentless" investigation tool that pushes its' application software packet across network resources directly into the memory space of suspect systems. Live Wire is designed for either permanent installation to a corporate network or ad hoc investigation. EnCase Enterprise is intended for permanent corporate implementation; a special version called the Field

Intelligence Model (FIM) is available to Law Enforcement only for ad hoc investigation. Prodiscover IR by TechPathways performs collection and analysis tasks and may be utilized ad hoc or via in situ implementation.

Ad Hoc scripts and programs

Even with the plethora of pre-built tools available to the forensics examiner many circumstances may arise in which a custom solution may need to be created in order to fully address investigative needs. Code writing, scripting and resource fabrication skills are a welcome addition to the forensic examiner's toolbox! Of course, any ad hoc implementation still requires full validation and testing of mechanics and process in order to guarantee the tool performs as expected and preserves the integrity of the evidence upon which it acts.

Software Licensing

Quite simply, and examiner's entire software environment (encompassing operating system(s) and applications) must be license legal at all times. Utilizing unlicensed software is effectively and act of software piracy (illegal). Any evidence gathered during the commission of a crime may prove inadmissible. Software licensing is a main cost factor for laboratory scope of service provision, initial purchasing targets and ongoing laboratory operations. Maintaining, auditing, documenting and demonstrating license compliance should be integrated into the laboratory SOP.

Tool Validation

Tool testing is a mandatory and ongoing process for both a forensics examiner and a forensics laboratory. Both software and hardware need to be tested to insure they perform as expected. Repeat testing validates method and guarantees equipment integrity over time. It is critical to demonstrate and document proof of any given tool's ability to preserve the integrity of any data under examination. Testing and documentation of testing are the foundation of a very important line of process defensibility: an examiner needs to be able to say with certainty (i.e. provide test results) the tools utilized in evidence examination do not spoliate the evidence. No examiner wants to be on the witness stand and be unable to provide proof of the competency of their methodology and tools. The best way to prevent this unfortunate circumstance is to consistently and repeatedly test every hardware and software tool, document the test methodology, results, and theory surrounding the test design, and be able to state that testing was performed personally as well as departmentally. The best defense toward methodology attacks is to be able to state "I personally tested this tool/method/process. Here is a copy of my test results. Let me explain why my testing is valid proof of the soundness of my method. Would you care to see a demonstration?" The exact methodology of tool testing needs to be tailored to the tool. Testing must be specific

and exhaustive enough to prove the underlying function of the tool valid. For instance, most examiners will create a specific test script for their write block devices, create a "gold standard" test media, and perform a set number of procedures (attempt to write data to media, attempt to delete data from media, attempt to access media via X method, etc.) in order to demonstrate tool function integrity. The NIST (National Institute of Standards and Technology) has an entire web site dedicated to the testing of computer forensics tools (www.cftt.nist.gov/). Of note on the NIST website are a general framework for tool testing and a listing of several specific tool test documents.

NOTE

1. *Cyber Crime Investigations: Bridging the Gaps between Security Professionals, Law Enforcement, and Prosecutors* (Elsevier, 2007).

Chapter 2

Digital Forensics: An Overview

Solutions in this chapter:

- **Digital Forensic Principles**
- **Digital Environments**
- **Digital Forensic Methodologies**

- ☑ **Summary**
- ☑ **Solutions Fast Track**
- ☑ **Frequently Asked Questions**

Introduction

Through the efforts of computer scientists, law enforcement and intelligence officers, network and system administrators, programmers, academics, and hobbyists, the field of digital forensics has evolved (and is still evolving) into one of the most dynamic and powerful investigative techniques in use to date. Digital forensic managers and investigators face a host of procedural, legal, operational, and technical challenges driven by both the explosive pace of technological innovation and the sophistication of cybercriminals. To assist them, this chapter will provide an overview of digital forensic principles and methodologies and the differing digital environments encountered, and introduce concepts covered in greater detail throughout this book.

NOTE

There are numerous definitions of computer (digital) forensics; my favorite comes, in part, from Dan Farmer and Wietse Venema: "Gathering and analyzing data in a manner as free from distortion or bias as possible to reconstruct data or what has happened in the past on a system" *and providing clear and objective testimony and reporting of the results of the investigation*. (The italics represent my addition to Farmer's and Venema's definition.)

Digital Forensic Principles

Before we move into a discussion of digital forensic principles, it is important that we understand the difference between principles and procedures (methodologies). The Merriam-Webster online dictionary defines a *principle* as "a comprehensive and fundamental law, doctrine, assumption or rule" and a *procedure* as "a particular way of accomplishing something or of acting." The difference between the two terms can appear to be minimal, but it is important: A principle is a fundamental truth that governs a specific endeavor; in contrast, a procedure is a method of accomplishing something.

Practice Safe Forensics

Although it has little to do with the technical side of digital forensics, safety is one principle that cannot be waived. This principle is invoked primarily while acquiring digital evidence, but you also need to keep it in mind during routine operations in the forensic lab. Regardless of whether the digital evidence is being acquired in the course of a high-risk, dynamic-entry search warrant or an administrative search of an employee's computer, the digital investigator needs to perform a risk assessment.

A key component of safety is analyzing risk. A risk assessment can be a formal process as defined by your agency or employer, or it can be an informal analysis of the risks associated with your activity. Things to consider include:

- Will the individual(s) whom you are investigating be present?
- If present, do they have a propensity toward violence?
- Are there weapons in the area?
- Is the area in which you will be working unsafe (i.e., does it contain hazardous materials, heavy machinery, etc.)?
- What is the surrounding area like (office park, college campus, economically depressed, etc.)?
- Will you be able to isolate your work space from bystanders?
- What time of day are you conducting your operation?

This list is a starting point for the types of risk assessment questions that you need to answer before conducting operations outside the office. For digital forensics conducted in your office/lab, the preceding questions are obviously superfluous. In that environment, your concern is to ensure a safe, accident-free workplace.



WARNING

It's tempting to bypass performing a risk assessment when conducting civil or even white-collar digital forensic investigations. Don't fall into that trap. The subject(s) of your investigation may be experiencing an extreme amount of stress as they contemplate the loss of their job, the income that supports them and their family, and/or the embarrassment and loss of social status as a result of being fired or arrested, and may decide to strike out against you. Be polite, professional, nonconfrontational, and observant and always have a plan to deal with potential irrational actions.

Establish and Maintain a Chain of Custody

Whether you are performing digital forensics for a civil lawsuit, a criminal investigation, or an administrative inquiry, the authenticity and integrity of the evidence you examined will be of critical importance.

Your first step is to establish a chain of custody policy for your organization. The goal of the policy is to ensure that each piece of evidence collected is accountable to an individual until it is either returned to its original owner or disposed of. You can do this by having each

person who possesses the evidence, from the original collector to the digital forensic analyst, the evidence custodian, and finally, the original owner, sign a receipt for the item, creating a "chain." Delineating a complete chain of custody policy is beyond the scope of this chapter, but some points to consider include the following:

- Designate an evidence custodian and backup evidence custodian.
- Build a limited-access evidence storage facility.
- Log access and activity in the evidence storage facility.
- Establish a Standard Operating Procedure (SOP) for collecting, marking, transporting, and storing evidence.
- Create appropriate forms to include chain of custody forms, evidence tags, evidence logs:
  - Each piece of evidence should be marked with (at a minimum) the initials of the person collecting the evidence; a unique evidence number; the location, date, and time it was collected; and its overall condition (to include identifying numbers).
  - Chain of custody forms should contain (at a minimum) a description of the evidence item (to include identifying numbers); the location, date, and time it was collected; the unique evidence number; from whom the evidence was obtained; the name of the person who first collected it along with the reason for collection; and the names of all subsequent people who signed for the evidence along with the date and reason for their taking possession of the evidence, ending with the final disposition of the evidence.
  - Evidence logs track the evidence that enters and leaves the evidence storage facility and should contain (at a minimum) a description of the evidence item (to include identifying numbers); the unique evidence number; the date and time the item entered or left the evidence room; and the name of who signed it in or out of the evidence room along with the name and signature of the evidence custodian.
- Audit the evidence storage facility and evidence logs and have someone other than the evidence custodian account for all evidence on a regular basis.



WARNING

If your digital forensic analysis becomes involved in adversarial proceedings, regardless of whether they are civil, criminal, or administrative, an error or omission in your evidence management could lead to your analysis being excluded from consideration by the presiding authority. Not a happy circumstance.

It is not enough to establish only the authenticity of your digital evidence; you will also need to establish its integrity. Because digital evidence consists of easily altered magnetic, electronic, or optical signals that cannot be seen or heard until they have been interpreted by hardware and software, special care must be taken to prove that your evidence was not modified during its collection, analysis, or storage. The primary method for proving integrity is the calculation of a hash value for each piece of evidence before and after it has been analyzed.

The National Institute of Standards and Technology (NIST) defines a *hash code* (value) as a "large number computed from the entire string of bits that form the file. The hash code (value) is computed in such a way that if one bit in the file is changed, a completely different hash code (value) is produced. To minimize the possibility that two different files may generate the same hash code (value), a sufficiently large hash value is computed."

The general procedure for establishing the integrity of digital evidence is as follows:

- Calculate a hash value of the digital evidence.
- Conduct a digital forensic examination using the appropriate procedures.
- Calculate a second hash value of the digital evidence and compare it to the first value. If they match, the digital evidence has maintained its integrity.

NOTE

In 2004, a team of researchers was able to defeat the MD5 hashing algorithm by identifying a hash collision (two files that have the same hash), and in 2005, parts of the same team identified a hash collision in the SHA-1 hashing algorithm. Although these discoveries were extremely significant in the cryptographic community, their impact in the digital forensic community has been somewhat muted. Current research by AccessData and NIST takes the position that the MD5 and SHA-1 algorithms are still secure for use in establishing digital evidence integrity and conducting forensic analysis. For those practitioners that want to be on the safe side, evidence and files can be hashed with both MD5 and SHA-1 with the probability that two different files will have the same MD5 and SHA-1 hash value computationally unrealistic for the foreseeable future.

Minimize Interaction with Original Evidence

Hollywood has discovered digital forensics. In some ways, it's a little flattering to have people ask you whether you can do what they saw on "CSI" or "24". On the other hand, you then have to explain that contrary to the depiction of digital forensics in the movies and on TV, you do not immediately hop onto the keyboard of the subject's computer/cell phone/PDA

(after single-handedly beating four terrorists into unconsciousness), defeat the password/encryption scheme, immediately locate the critical information, and call back to Chloe at the CTU.

In the real world, every digital forensic best practice I've ever reviewed stresses the criticality of not modifying the original evidence and creating a verified, bitstream image of the original evidence. Because the Federal Rules of Evidence allow the use of bitstream images to be considered as the "best evidence," most standard operating procedures for creating a forensically sound bitstream image include the following steps:

1. Sanitize your evidence storage devices by wiping them at the physical level and then partitioning and installing the appropriate file system.



TIP

Wiping a drive refers to overwriting each addressable sector with either a random or a discrete character. Numerous utilities are available for wiping hard drives; a search on Google using the term "wipe drive" will return numerous hits, both proprietary and open source. The Department of Defense (DoD) National Industrial Security Program Operating Manual 5220.22-M's clearing and sanitization matrix and NIST's Guidelines for Media Sanitization (NIST Special Publication 800-88) are excellent references.

2. Observe and record the physical characteristics of the digital evidence (location, cabling, peripherals, and status).
3. If necessary, use trusted software tools to collect volatile information (memory contents, network connections, etc.).
4. Disconnect the computer from any network connection (modem, Ethernet, wireless) if possible.
5. If the computer is powered on, note what is displayed on the screen (a digital camera is very useful here).
6. Power off the computer by pulling the power cable from the computer.
7. Either remove the hard drive(s) from the computer and attach them to a hardware write blocker and connect them to the imaging computer/device, or attach a hardware write blocker between the hard drive(s) and the data cable on the subject's computer.
8. Attach the appropriate digital storage device to either the imaging computer/device or the subject's computer.
9. Boot the imaging computer/device and obtain physical characteristics from the digital evidence drive controller (capacity, firmware version, geometry, etc.).



WARNING

If your digital evidence is an IDE or Serial ATA hard drive, be sure to look for either a host-protected area (HPA) or a device configuration overlay (DCO). Both methods were implemented by hard drive manufacturers to "hide" a portion of the hard drive from the operating system. The standard reason was to provide computer OEMs such as Dell, Sony, and HP a place to store the files necessary to restore the operating system. The HPA and DCO can be accessed by third-party tools to store data that is effectively invisible to the operating system.

10. Follow the protocol of computing a hash value for the evidence, imaging the evidence, computing a hash value for the evidence image, and comparing the hash values for the evidence and evidence image.
11. Copy the evidence image to a backup storage device before performing any analysis, and store the original evidence image in a controlled evidence storage facility.
12. Create and maintain the appropriate documentation.

It's important to stress that although the preceding protocol constitutes a sound method of collecting digital evidence for analysis, sometimes you may have to skip or violate some of the steps in the course of your investigation. Doing so will not automatically invalidate the digital evidence or analysis, *but* your rationale for the actions you did or didn't perform will be closely scrutinized. All I can offer by way of advice is that if you find yourself in that situation, document everything and be prepared to justify and support your actions.

Use Proven Tools and Know How They Work

In short, the tools make the digital investigator. Users interact with their digital devices through software, and that interaction creates the information that digital investigators collect, analyze, and interpret. The only way to preserve, collect, analyze, and interpret the data on digital devices is by using software (both standard and specialized). Another layer of complexity is that the findings of digital investigations are usually presented in administrative, civil, or criminal proceedings and their introduction and applicability are governed by rules of evidence. This makes selecting the appropriate hardware and software tools to accomplish a digital forensic task critically important.

Is my revision to the first sentence in the preceding paragraph OK?Luckily, although the discipline of digital forensics is constantly evolving, it is a mature field with a sizable and

active body of research, numerous hardware and software tools, and two or more generations of digital investigators who have contributed through trial and error to a substantial body of practical knowledge. In addition, the field of digital forensics is also extremely popular, lucrative, and competitive, so current hardware and software are rapidly updated and new hardware and software tools are constantly being released.

Therefore, instead of extensively searching to find one appropriate hardware or software tool, the digital investigator usually locates several. In an ideal world, a digital investigator would acquire and use all of them. However, digital investigations and investigators rarely exist in an ideal world, and temporal and financial constraints will usually limit the number of tools selected for the forensic toolbox. In the following sections, we will discuss some considerations in the selection process.

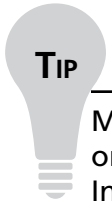
Is the Tool in General Use?

This is an important consideration because the U.S. legal system is predisposed to accept the results generated from standard, generally available hardware or software as opposed to custom hardware or software that is not in widespread use. This is not to imply that you should shy away from hardware and software from smaller developers; on the contrary, some very good tools come from the "little guys" (one example is Steve Gibson's Automated Image and Restore, http://air-imager.sourceforge.net). However, it does mean that if you decide that Tool X from a small developer is the best tool for a task, you have to make sure you conduct the appropriate due diligence before you use it. There is no one appropriate due diligence procedure, but you may want to consider some of the following paragraphs.

What Is the History of the Developer and the Tool?

Unless you are using a hardware or software tool that has just been developed (not an ideal situation, but one that may come up), your particular tool will have a history behind it. That history will include the background of the developers, who they are, what have they developed previously, their nationality (this may have some bearing if you work on government contracts), their political philosophy, and similar information. The majority of this information should be readily accessible via targeted Internet searches, and it goes without saying that most of us leave pretty significant digital tracks.

Additionally, if your particular tool has been in use for a period of time, it will have attracted proponents and opponents, and they are usually quite vocal. A Google search on your tool should turn up a reasonable amount of information from publicly available resources. For example, NIST has conducted publicly available research on disk imaging and hardware and software write blocking, and Yahoo! Groups hosts several forensics-related mailing lists.



TIP

Most digital forensic practitioners are members of one or more professional organizations, among them the International Association of Computer Investigative Specialists (IACIS; www.iacis.com), the International High Technology Crime Investigation Association (HTCIA; www.htcia.org), and the International Society of Forensic Computer Examiners (ISFCE; www.isfce.com). These organizations maintain members-only mailing lists and discussion forums which are excellent sources of information and opinions regarding digital forensics. The membership fees are reasonable, and you must meet certain entrance requirements (some are law enforcement only). I recommend joining one or more groups.

Perhaps the most important consideration pertaining to the history of a tool is how it has fared in the legal arena. Some tools have been frequently involved in civil and/or criminal litigation and have extensive records. Accessing those records without access to LexisNexis or Westlaw may be difficult; however, it may pay to contact the developer and ask whether they have a record of the litigation their product has been involved in. For example, Guidance Software does an excellent job in this regard with its EnCase Legal Journal. It goes without saying that if the accuracy or reliability of your preferred tool has been discredited in a civil or criminal legal proceeding, continued use will be problematic, to say the least.

Do You Know How the Tool Works?

A thorough understanding of how their equipment functions is absolutely critical for digital forensic examiners. Luckily (in the United States), the courts do not require that digital investigators demonstrate an extremely granular knowledge of how a particular piece of hardware or software performs. This doctrine does *not* release the digital investigator from demonstrating that he understands how the tool works at a relatively high level and displaying competence in operating the tool. Successful technical and legal challenges to the accuracy and reliability of tools are relatively rare, *but* attacks on the digital investigator's knowledge and competence are very common. These attacks manifest themselves during the testimony of a digital investigator in the course of either an evidentiary hearing or litigation, and may result in the exclusion of the evidence or a reduction in the weight of the evidence.

To avoid those unpleasant outcomes, digital investigators needs to become extremely familiar with the basic theory and practice of digital forensics, as well as the actual operation of their tool. There is no one "right" way to obtain this knowledge and experience. I have worked with both self-taught and formally trained digital investigators who were extremely competent and professional. For those who pursue the self-taught approach, a wealth of

information is available at bookstores as well as online. Meanwhile, formal training for digital forensics has exploded; when I began my career, the only formal training was available from the government in law enforcement, the intelligence community, or the military. Those programs continue to turn out highly skilled digital investigators, but in addition to the government, a digital investigator can begin his career in academia by earning a degree in computer forensics or through one of the many private sector digital forensic training programs, both public and in-house.

A combination of training and experience is the best way to obtain a thorough familiarity with a piece of hardware or software. Either vendor or third-party training is perfectly acceptable, but there is no substitute for experience. That experience is gained by using the tool in a controlled environment to confirm that it is reliable and accurate while increasing your understanding of its functionality. This applies to standard software (robocopy.exe, dd, grep, etc.) as well as specialized software (SMART, EnCase, the Sleuth Kit, and so on).

A thorough understanding of how a tool works and as well as its functionality is important—not only to the eventual proceedings for which you are collecting and analyzing evidence, but also when Murphy's Law is invoked. Murphy's Law states that what can go wrong will go wrong, and it will happen at the worst possible time, usually while you're collecting digital evidence or conducting onsite analysis and you hare forced to overcome a problem (software gremlins, hardware failures, etc.) with limited resources (time, software, hardware, etc.). Knowing how to leverage your resources can mean the difference between success and failure.

Conduct Objective Analysis and Reporting

Because of the complexity of the technologies that support our society and the general aura of "wizardry" surrounding digital forensics, the importance of being an impartial fact finder cannot be overstated. The consumers of our reports, regardless of whether they are in the administrative, civil, or criminal arena, probably won't be conversant on the topic of digital forensics. They will be relying on you, the digital investigator, to explain and interpret the information contained in their reports. This requirement to provide explanations and interpretation conveys the potential to significantly influence the outcome of a proceeding. As such, digital investigators must take care to ensure that their analysis is both impartial and thoroughly researched.

Analysis of the information identified during a digital forensic examination can be a slippery slope. For example, through an analysis of files, their metadata, and system logs, the creation, copying, and deletion of a file can be identified. A digital investigator can testify that these actions occurred, but should not testify (based on only the evidence on the computer) as to the mindset of the individual who performed them.

Most digital investigators are technophiles (myself included) who are thoroughly engaged in the intricacies of digital forensics; however, one of the most critical parts of our job is to

communicate the results of our analysis in a manner that can be easily comprehended. The technical expertise we display in the collection and analysis of digital evidence is meaningless if no one can read or understand our reports. This is not to disparage the many lawyers and judges who intimately understand digital forensics, but one of the greatest challenges digital investigators face is to simplify their reports so that they can be used in the legal system while adequately conveying the necessary technical information.

Digital Environments

The tools that digital investigators use are incredibly varied, but in the final analysis, they all just process and store information. The environments they are used in are significantly more varied, and will likely present greater challenges to the conduct of digital forensics than will technical issues.

Corporate

Calvin Coolidge was right about at least one thing: "The business of America is business." As such, most digital forensic investigation will at some point interact with the corporate world. For clarity's sake, the *corporate world* refers to large businesses (Fortune 1000) rather than smaller entrepreneurs. The information infrastructure in the corporate world is characterized by dedicated IT departments with policy and operations managed by a chief information officer (CIO), a centralized help desk, systems and network administration, a modern, hardware and software, disaster recovery protocols with off-site data storage, a data backup policy that is (usually) followed, and a dedicated IT security staff. The businesses in the corporate world have legal departments (or law firms on retainer) that are experienced in electronic discovery, and (in the cases where they aren't the subject) are cooperative with digital investigators. Obtaining this cooperation is the challenge; unless your investigation involves exigent circumstances, businesses will generally require a court order, subpoena, or search warrant before they allow the release or collection of information.

Government

Digital investigators need to avoid the misconception the government information infrastructure is standardized. In fact, a mind-boggling variability exists within federal, state, and local governments. Some agencies closely resemble corporate world information infrastructures, whereas other agencies operate with extremely limited resources and lack robust management. Although federal agencies usually have greater resources to devote to IT infrastructure than state or local agencies, that is not always the case. Local governments in wealthy communities will frequently have superior IT resources than the federal or state regional office in that area. Most government agencies will have either an IT security staff or a relationship with the appropriate law enforcement agency so that they can perform their

own administrative investigations and collect evidence for criminal investigations. Conducting an official investigation in this environment will usually be a relatively straightforward affair. Private sector investigations in the government environment are somewhat rare and will certainly require the use of a court order or subpoena to secure the release or collection of evidence.

Academic

The academic environment encompasses all levels of educational and research institutions, from kindergarten to supercomputing centers, and the IT architectures will be similarly varied. Investigators can expect to encounter somewhat chaotic information infrastructures and a lack of centralized management. The level of hardware and software can vary from archaic but functional, to bleeding-edge research. The level of cooperation will vary greatly, and investigators will need to be prepared to obtain court orders, subpoenas, or search warrants for all but the least intrusive investigative techniques.

The Internet

Although the Internet doesn't exactly resemble what William Gibson and Bruce Sterling envisioned in *Neuromancer* and *Snow Crash*, it's close, and in spite of the efforts of governments, corporations, and lawyers, it is still a relatively open and anonymous world (in general, a good thing) that unfortunately offers a refuge for a shadowy, transnational community of gray market Internet service providers (ISPs), darknets, fraudulent Web sites, and botnets. Investigations in the Internet environment will be challenging and rely heavily on the ability of the investigator to obtain either cooperation or legal tools to collect evidence.

The Home

The home environment encompasses everything from grandma's computer to a small-business network. These environments will be greatly determined by the abilities and interests of the owner. A digital media enthusiast or hard-core computer gamer will have a radically different environment than an entrepreneur running a small business. In the first example, you can expect to encounter high-end digital devices, data storage, a fast Internet connection, and a home network (wired and wireless). In the second example, you can expect to encounter the bare minimum of IT resources to keep the business running.

Digital Forensic Methodologies

Karl Von Clausewitz was talking about war in the 19th century when he said, "Everything in (war) is very simple. But the simplest thing is very difficult." However, his analysis can easily apply to digital forensics in the 21st century.

Back in the "good old days" (the mid-1990s), there was computer forensics, and everyone generally understood that when you performed computer forensics you examined hard drives, floppy disks, CD-ROMS, and so on. The coming of the Internet Age ushered in online investigations, e-mail, Web site defacement, and the tracking of "hackers." Now we're in the Information Age where everything and everyone is online and wireless, every investigation has a digital component, and everyone is doing digital forensics in one way or another. To provide some clarity, I've divided digital forensics into three categories. Although they all deal with digital evidence, the manner in which they are approached varies.

Litigation Support

Litigation support is both a distinct category in the legal field and a category of digital forensics. To differentiate, we will stipulate that anything dealing with paper-based information or evidence is in the purview of litigation support as a legal discipline. In the context of digital forensics, litigation support entails the identification, collection, organization, and presentation of immense amounts of digital data. In practice, this occurs during civil litigation and U.S. Securities and Exchange Commission (SEC) enforcement actions, but if recent white-collar prosecutions are any guide, litigation support is becoming prevalent in the criminal realm as well.



WARNING

In general, during litigation support, evidence destruction or alteration is not an overriding concern, but you need to be aware of the possibility. If you suspect that evidence destruction or alteration has been committed by either party, you have a duty to inform the attorney for whom you are performing litigation support. You can be certain that an assertion of evidence destruction or alteration will be controversial; however, you cannot let that deter you. Double-check your findings and *document everything*!

Identification

The identification phase entails analyzing the information infrastructure of the opposing party and advising the lead attorney regarding the location(s) of evidence that may be relevant and how best to collect it. You will also perform this analysis on your client's information infrastructure so that you can respond to the opposing party's requests for information. All cases are different, but I've never heard of a case that didn't involve identifying and locating the e-mail (now instant messages and text messages), documents, and spreadsheets of the key people.

Collection

The next phase involves collecting the digital evidence through the use of civil subpoenas, discovery orders, and so on. When responding to subpoenas and discovery orders, remember that they have deadlines for the production of evidence and failure to meet these deadlines can result in sanctions. When you are collecting evidence from the opposing party you will either receive what was listed on your subpoenas and discovery orders, or be allowed to collect the information personally. Planning is critical; you may get only "one bite at the apple," and if you are unable to collect the information, you may not get another opportunity.

Tools & Traps…

It's Still All about People!

In spite of the perception that digital forensics is solely a technological endeavor, interpersonal skills are critical for the successful digital investigator. You may be wondering why interpersonal skills are an important tool or why I'm mentioning them in the context of litigation support, especially because attorneys conduct the majority of the interviews and depositions. The answer is that while performing litigation support, you will come into contact with the IT staff of the opposing party, and (unsurprisingly) they won't be predisposed to be helpful. It will be up to you to build the rapport necessary to establish a working relationship that will gain their confidence and cooperation. Regardless of whether you are dealing with a network administrator or the only guy who knows how to run the tape robot, you all have a common denominator of working in the IT field, and probably have similar experiences and interests. You would be surprised how many times I have had colleagues tell me that just when they thought they had collected all the information they could find, someone on the IT staff happened to mention that "box of old e-mail backup tapes from 2002 that are on the back shelf of the closet…"

Organization

Once you've collected all your information—and in today's world, we're probably talking about terabytes (TB), not gigabytes (GB)—you've got to organize it so that the attorneys can review it. This usually means indexing all the digital evidence and conducting keyword searches, and then supplying those results to the attorneys. This is generally an industrial task in which you balance the index and search algorithms of your software and the computational power of your hardware against the selectiveness of your search criteria

(keywords, date ranges, user activity, etc.) and the size of your data. Several commercially available, specialized indexing and searching programs can help with this (I'm partial to dtSearch), as can desktop search solutions (from Google, Apple, and Microsoft, among others) and, of course, today's powerful open source utilities (e.g., grep, find, PERL, etc.). Because acquiring computational power and data storage has become a matter of budgetary constraints only (that is, until we routinely start dealing with petabytes), your primary challenge will lie in working with your attorneys to create the most selective search criteria possible. In my experience, attorneys tend to err on the side of expansiveness, and as their technology advisor, you can best assist them in leveraging their search. If your case is particularly complex or involves a large number of entities (people, corporations, bank accounts, etc.), it may be necessary to import the data into a data visualization package such as Analyst's Notebook.

Presentation

Presentation has two components. The initial one is the process whereby you transmit the results of the organization phase to the attorneys for their review; the second is the presentation of selected evidence in a legal or administrative proceeding. You may be tempted to think that once the process of indexing and searching the digital evidence is completed, the hard work of litigation support is over, but you would be wrong. The hard work is actually just beginning. Even with the most targeted search criteria, usually an immense number of evidence items will need to be manually reviewed for relevance and privilege claims. As the digital investigator, it will be your responsibility to implement a system that will allow the legal team to review evidence, extract relevant items, segregate potential privileged materials, and conduct additional searches, while maintaining the integrity of the evidence and workflow management. This can be as simple as providing each attorney with a DVDR containing a subset of the evidence to review with a desktop search program to provide a local search capability. Items that are considered relevant or privileged can be copied to directories on a shared network drive for additional review. This approach can be quickly and simply implemented and will work extremely well for less complicated cases. It does not, however, scale up very well, it offers little in the way of collaborative workflow, and it has the potential to be very manpower-intensive. For gargantuan cases (Enron, Tyco, etc.), your options devolve to either building your own infrastructure or subcontracting to a specialist such as KPMG's Forensic Technology Service. We're not going to discuss trial presentation; although you'll have input and be a key part in preparing the presentation, the overall direction and strategy will be determined by the litigating attorney(s).

Digital Media Analysis

As mentioned earlier, the discipline of computer forensics was involved in collecting and analyzing digital evidence from computers and their associated removable media. However, computers and digital storage have evolved significantly in both form and functionality, and

together they have proliferated into every facet of our personal and professional lives. Although the term *computer forensics* is semantically still an accurate way to conceptualize the collection and analysis of digital evidence, it is limited by society's dated vision of a computer as a beige box that sits on a desk, when in reality computers are everywhere and in everything. The term *digital media analysis* is an attempt to convey the expanded possibilities for digital evidence collection and analysis that are present in the Information Age.

Media analysis is a component of both the litigation support and the network investigations categories, but it is also a distinct category of digital forensics. The overwhelming majority of all investigations (civil, criminal, or administrative) involves digital evidence with digital investigators performing media analysis in cases ranging from employee misconduct to murder investigations. This breadth of exposure makes the journeyman digital investigator a valuable asset to the traditional investigator who is working the case, but perhaps the most important contribution has nothing to do with the collection and examination of digital evidence.

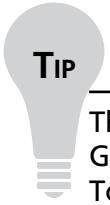
Identification

Although most traditional investigators understand that digital evidence can be valuable to their cases, their primary focus is on witnesses, interviewing, interrogation, and physical evidence. This is understandable, as these are important components of a successful investigation, but a digital investigator can identify sources of digital evidence that will enhance the effectiveness of traditional investigative procedures. Most traditional investigators will consider seizing the subject's computer and e-mail and obtaining toll records, but may not consider some of the following digital evidence:

- Instant messages (IMs) and Short Message Service (SMS) messages
- Online storage, including dedicated (Apple's .mac accounts) and unintentional (Google's GMail)
- Multiple e-mail accounts
- Cell phone video, data, and pictures
- Facebook, MySpace, personal Web sites, forum postings, and blogs
- iPods, digital video recorders (DVRs), and personal media players (PMPs)
- USB thumb drives (in all conceivable shapes and sizes)
- Wireless network storage
- Internet artifacts (cookies, Internet cache, auto-complete entries, etc.)

The preceding list is a starting point in the search for digital evidence, but information developed from sources on the list has proven extremely useful as stand-alone evidence and

in the corroboration and impeachment of witness and subject statements. A key consideration is the timely identification of digital evidence, that is, some of the online evidence may not be maintained by the network provider past a certain point in time (usually 30–90 days), and the digital storage on mobile devices is subject to loss if the power is disrupted. As soon as relevant online evidence has been identified, collection and analysis planning needs to begin. This does not imply that the identification phase is complete; digital investigators need to remain constantly alert for new and innovative types of digital evidence.



TIP

This is an excellent justification for checking Engadget (www.engadget.com), Gizmodo (www.gizmodo.com), Ars Technica (www.arstechnica.com), and Tom's Hardware (www.tomshardware.com) on company time!

Collection

Once potential digital evidence has been identified through the collaboration between traditional and digital investigators, the collection phase can begin. In contrast to litigation support, the collection phase of media analysis is a more adversarial process without pre-collection meetings to establish the breadth and format of the digital evidence, and it generally involves taking possession of digital devices to either image them or seize them as evidence under some type of civil subpoena, administrative order, or search warrant.

Preparation and planning are therefore critical to success. Hopefully during the course of the investigation, you will have developed information regarding the digital devices, data storage, and communication practices of the target(s) of your investigation. In an ideal world, you will be able to locate a technology-savvy cooperating witness, and in white-collar, financial crime cases, it is entirely feasible that someone will be willing to provide that type of information. Unfortunately, in non-white-collar investigations, you will have a much more difficult time locating that type of cooperating witness. Regardless of the quality of your cooperating witnesses, you can be certain that some of the information they provide will be incorrect and out of date and you will need to factor that into your preoperational planning. The collection plan will need to address the following issues:

- Safety and security of the collection scene
- The subject(s), violation/allegation, and a brief synopsis of the investigation
- A description of the location
- The digital evidence authorized to be collected and under what authority
- The digital devices and data storage assumed to be present and where they are located

- The priority for collection and personnel assignments
- The method of digital evidence collection to include equipment, data format, and documentation
- Evidence marking, handling, and control
- Chain of command and contact information

Due to the inherently technical nature of digital forensics, equipment preparation will be a significant factor in the success or failure of digital evidence collection. Because there will always be uncertainty as to what you will encounter when you serve your subpoena or execute your search warrant, the Boy Scout motto of "Be Prepared" is certainly appropriate. However, unless you have an unlimited budget, you can't take your entire digital forensic lab to every collection site, so you've got to make choices based on the information you have at the time. I wish I could provide a foolproof guide to selecting the correct hardware and software, but because I've never been on or met anyone who has been on the same collection operation twice, I can only provide the following recommendations:

- Always bring a flashlight (check the batteries) and toolkit with security bits for the screwdriver.
- Bring as many drive adapters, cable adapters, and write blockers (don't forget the cables and power adapters) as you can fit in your search kit.
- You can never go wrong with one more boot CD (don't forget to consider non-x86 architectures).
- Bring at least one spare FireWire/USB PCI card and FireWire PCMCIA or ExpressCard.
- Bring at least one more spare hard drive than you think you need.
- Bring a small 10/100 (Gig-E, if you can) switch or hub and Ethernet cables.
- Someone on the team will need to bring a cell phone/mobile device seizure kit.
- A multitool (Leatherman, Gerber, etc.) is a critical piece of equipment.
- Bring at least one extra power strip.
- Bring something that can detect wireless networks (laptop, Wi-Fi finder, etc.).
- Bring a small notebook and pens and take copious notes. The scribbled note you wrote at 3:00 A.M. may be important.

The preceding recommendations are derived from my firm belief and experience that although you are on a digital evidence collection operation, something will break or not work and you will need to be prepared to troubleshoot and solve problems. The items in this

list are in addition to whatever hardware and software you determine you will need to address the digital devices and data storage you plan to encounter. It may make a great story to say that you collected all the evidence using nothing more than a Leatherman Tool, a USB hard drive, and a Helix boot CD, but I wouldn't recommend that as a standard practice.

Once you are on the scene, your actions will be governed by digital forensic principles and the digital evidence collection methodology of your client or employer. If you need a reference for digital evidence collection, the United States Secret Service provides a best practices guide at www.secretservice.gov/electronic_evidence.shtml.



WARNING

In the course of collecting digital evidence, you may need to ignore your client's or employer's evidence collection methodology due to time issues, technical problems, and so forth. My advice in these inevitable situations is to document everything, and then determine whether your course of action will violate any digital forensic principles. If your course of action will not violate any principles, explain the problem and your solution to your client or on-scene supervisor. Once you have a decision, document it and continue collecting evidence.

Analysis

Analyzing digital evidence may not be as exciting as identifying and collecting it, but it is the most critical component of media analysis. In this phase, you extract data and interpret artifacts to create a report that organizes and interprets the arcane world of digital evidence so that it can be used to prove or disprove civil, administrative, or criminal allegations. The dizzying array of digital devices and data storage media (servers, network storage, RAID arrays, Storage Area Networks (SANs), iPods, USB thumb drives, cell phones and mobile devices, etc.) can be simplified into hardware and file systems. If the operating system and hardware on your digital forensic workstation can interact with the digital device you are analyzing through the use of adapters (cell phones are particularly finicky), drivers, or (if necessary) an abstraction layer (a hardware RAID array that presents itself as one unified drive), and your forensic tool can read the file system, you will be able to conduct an analysis. If you're lucky, most of the compatibility issues were addressed during the collection phase and you are merely dealing with image files.

The analysis of digital devices and storage to locate and extract data that is relevant and has stand-alone probative value has evolved into an advanced methodology that is driven by development of ever more powerful and sophisticated digital forensic tools. The introduction (and ongoing development) of automated, inclusive digital forensic suites with graphical user

interfaces such as iLook, EnCase, and FTK (and others!) has transformed media analysis. Instead of using several specialized programs to perform media analysis, you can use one program and accomplish most of these common tasks:

- Identifying files by their headers and grouping them by type
- Recovering deleted files
- Searching files and slack/unallocated space
- Extracting and processing e-mail
- Parsing log files and analyzing the Windows Registry
- Performing metadata and time line analysis
- Creating reports

In addition, some tasks, such as e-mail processing and file undeletion, are performed automatically and in the background. Automated, inclusive digital forensic suites are the rule for media analysis and have been a tremendous productivity multiplier. This is not to denigrate the universe of command-line forensic tools in any way! Some command-line tools do an outstanding job of media analysis, including Foremost, the Sleuth Kit, and the gamut of open source utilities currently available. Assuming that the user of the digital evidence you are analyzing has not employed any advanced antiforensic/data-hiding techniques (discussed shortly), an experienced digital forensic examiner is going to be able to extract all relevant information that has stand-alone probative value using an automated, inclusive forensic suite or command-line utilities. In fact, without a precise keyword list or date/time range, you will run the danger of producing too much information and overload the case agent.

Operating system artifacts is a term to describe the data (metadata, log files, inodes, plists, restore points, temporary files and directories, etc.) that all operating systems create as they perform the myriad functions necessary to run a digital device. These artifacts are only rarely apparent to the casual user, but when identified, extracted, and (more importantly) interpreted, they can be used to re-create actions and states of digital devices. Although the dominant operating environment today is Microsoft Windows, which is proprietary with limited documentation, digital investigators can reasonably expect to encounter Linux, Mac OS X, and mobile device operating systems (Symbian and Palm, etc.) also. Although the digital forensic community is always uncovering more information on non-open source operating systems, artifact identification and interpretation will be more challenging than with open source (Linux, etc.) or partially open source (Mac OS X) operating systems. Once artifacts are located, the challenge for the digital investigator is twofold. First you must correctly interpret what information the artifact conveys, bearing in mind that different operating systems (and versions) treat ostensibly the same type of artifact differently (e.g., Linux and Windows have different definitions of Modify, Access, and Create times).

The second challenge is perhaps more daunting. You will have to explain to an audience of laypeople how the inner workings of an operating system created this very obscure bit of data, what the data means, and then (if their eyes haven't gazed over) why this is important.

Notes from the Underground…

Anti-forensics

As digital forensics has evolved and grown as an investigative discipline, the techniques and tools of *anti-forensics* have also evolved. Anti-forensics can be defined as both a body of research and the software tools intended to defeat digital forensic techniques and tools by either hiding information or challenging the presumption of reliability that mature digital forensic tools enjoy. Anti-forensic developers have created software to implement the following techniques in an attempt to defeat digital forensic analysis: hiding data in slack/unallocated space, altering file signatures, altering timestamps, secure deletion (overwriting), and running programs in memory only. Counters to most, if not all, anti-forensic techniques exist, such as performing a live response to collect the contents of the memory on a digital device. The Metasploit Project is an excellent place to begin researching anti-forensics (www.metasploit.com/projects/antiforensics/).

The final step is in the analysis phase is usually the least popular for digital forensic examiners, and I have *never* met one (myself included) who enjoyed writing the analysis report. It is, however, critical, no matter how brilliant and insightful your digital forensic examination was. If you cannot communicate your findings, your efforts will be for naught. Most employers and government agencies will have a standard report format, either a freeform narrative or a predefined form. In either case, you will need to provide some method to distribute the data you've extracted, usually a DVDR or (if you've got the secure bandwidth) a network share. My recommendation is to provide a brief freeform narrative report that reviews the collection and analysis processes and explains significant findings in enough detail for use in official proceedings, and then attach a DVDR that contains an in-depth report generated by the forensic tool suite used to analyze the evidence.

Network Investigations

As we discussed earlier in this chapter, the pervasiveness of digital devices in society almost guarantees that traditional investigations will have a digital evidence component with varying degrees of importance to the case. In contrast, network investigations are focused on collecting and analyzing digital evidence regarding the use of networked digital devices in

commission of crimes and unauthorized activities. Your objective is to determine whether a violation occurred, what happened, who did it, and where they are. Your investigation will begin with the victim digital device(s) and proceed through cyberspace to locate the attacker in the real world while obtaining enough evidence to support an official proceeding.

Identification

In network investigations, digital evidence comes in three forms: evidence of the crime, network evidence that allows you to trace the attacker, and personally identifiable information that will allow you to identify the attacker. This information will belong to numerous entities, be stored on varying digital devices, and can be physically located in multiple jurisdictions. Table 2.1 shows some examples of entities, digital devices, and potential evidence.

Table 2.1 Entities, Digital Devices, and Potential Evidence

Entity	Devices	Potential Evidence
The victim digital device(s)	Anything connected to a network that can store and process information	Evidence of the crime, trace information, and identifying information
ISPs	Routers, Internet Protocol (IP) address servers, customer/billing databases, network storage, etc.	Evidence of the crime, trace information, and identifying information to include billing information
Domain Name Service (DNS) registrars	Contact/Customer/Billing databases	Identifying information to include billing information
Co-location or Web hosting providers	Routers, firewall and intrusion detection system (IDS) logs, application server logs, network storage, and customer/billing databases	Evidence of the crime, trace information, and identifying information to include billing information
E-mail providers	E-mail servers with connection logs and stored e-mail and customer/billing databases	Evidence of the crime, trace information, and identifying information to include billing information
Internet Relay Chat (IRC) and Web forums	IRC server and connection logs and user registration databases	Evidence of the crime, trace information, and identifying information
The attacker's digital device(s)	Anything connected to a network that can store and process information	Evidence of the crime, trace information, and identifying information

At first, the preceding examples appear to provide a wealth of evidence. Unfortunately, appearances can be deceiving. Entities may not log connection information, the information they do collect will be maintained for only a finite duration due to limited data storage and data retention policies, identifying information may be falsified, the financial instruments used to pay for services may be stolen or anonymous (e-gold), evidence of the crime may be deleted, and the attacker may take steps to relocate. These concerns make the evidence identification and collection cycle extremely time-sensitive, and it may be necessary to triage network investigations so that the ones with the highest probability of locating and identifying an attacker receive the most resources.

Collection

As you trace the attacker to his point of origin through the Internet, you will locate new entities that possess digital devices containing potential evidence. Each time you locate one, you will collect and analyze whatever evidence you can obtain to identify the next entity on the trail leading back to the attacker. As you prepare to communicate with each entity, your will need to address the following questions:

- Is the entity reputable? (That is, will they cooperate or hinder and disclose your investigation?)
- Under what authority are you requesting or compelling cooperation?
- Does the entity have the technical expertise to preserve and extract digital evidence?
- Is their involvement a result of their normal operations (ISP, Web hosting, e-mail, etc.)?
- Is the entity also a victim of cybercrime (e.g.,, computer/network intrusion, theft of service, etc.)?

The first question is problematic. For questionable entities, some judicious "Googling" and trusting your instincts will have to suffice to resolve their legitimacy. Your approach will be further modified (or constrained) by the authority under which you are requesting or compelling cooperation. Law enforcement officers have the option of requesting consent to obtain evidence or using the tools (subpoenas, court orders, and search warrants) provided by the legal system to compel cooperation. Private sector digital investigators may also request consent to obtain evidence and may avail themselves of a subset (subpoenas and court orders) of the tools provided by the legal system. As a practical matter, U.S. privacy protection laws, including the Electronic Communications Privacy Act (ECPA) and the Right to Financial Privacy Act (RFPA), as well as concern over liability, will limit cooperation sans a subpoena or court order. The virtual nature of the Internet practically guarantees that you will encounter an entity located in a foreign country. As a law enforcement officer, you may

be able to work with the Department of Justice Computer Crimes and Intellectual Property Section to use Mutual Legal Assistance Treaties or Mutual Assistance Agreements to obtain evidence. As a private sector digital investigator, you may certainly request consent from a foreign entity, but to compel assistance your only option may be a *letter rogatory* which is the process whereby the courts in one country will honor the requests of the courts in another country as a matter of comity.

Assuming that you have a reputable entity that possesses potential evidence and will cooperate, your collection methodology will hinge on the competence of the IT security staff and whether the evidence is a result of their normal business or whether they are a victim. Most ISPs, network storage and e-mail providers, and Web hosting/collocation providers have dealt with providing evidence to digital investigators and have an established protocol in place. In those cases, you will merely need to provide the parameters (IP addresses, date and time ranges, server logs, data, etc.) of the evidence you have the authority to obtain and they will collect it and provide it to you (usually via DVDR or FTP). Even if the entities are the victims of cybercrime, they usually have the expertise in-house to investigate it and will still be able to provide relevant evidence. In cases where the entity does not have that level of expertise in-house, you will need to collect the evidence.

Because important pieces of information (open sockets, RAM, established sessions, etc.) exist only in memory, it is necessary to conduct a live response to collect that volatile evidence. *Live response* is a methodology that calls for using trusted software (usually from a custom live response CD, thumb drive, or network share) and remote data storage (network share, USB drive, netcat socket) to collect volatile information which would be lost when a digital device is powered off. Trusted software is used to protect against the possibility that the victim operating system has been compromised and that modified versions of system utilities have been installed. Remote storage is used to minimize modifications to the data storage of the victim device. Live response is a departure from the principle of minimizing interaction with the original evidence and requires a balancing of the potential evidence collected versus a compromise of evidence integrity. It should go without saying that live response is an advanced digital forensic technique and should be performed only by a properly trained and equipped digital investigator.

NOTE

Live response is a technique that has been mostly applied to network investigations; its use should become more prevalent due to a combination of anti-forensics, whole-disk encryption, and the inexorable growth of data storage. Many live response CDs are available, and one of the most popular is Helix (www.e-fense.com/helix).

Analysis

The avenues of analysis in network investigations include the. traditional investigative steps that are applied to customer, registration, and financial evidence collected to build a profile of the subject(s) of the investigation. The network trace evidence analysis will have been an ongoing process as your investigation tracked the attacker from one entity to another. That analysis can be complicated by the sheer volume of data (connection logs, router logs, application server logs, firewall and IDS logs, etc.), but the use of a visualization tool such as Analyst's Notebook or jpcap can help simplify the process. Issues that may prove problematic are time (accuracy, time zones, clock drift, etc.), dynamically assigned IP addresses, public Internet access points, and wireless access points (free and unsecured). Due to a combination of limited storage space, encryption, and the legal protections involving electronic communications, it is improbable that your network trace evidence will include content, though. Media analysis bookends your network investigation. In the beginning, you will have conducted an analysis of the victim digital device(s), and once you have located and identified the attacker(s), you will (hopefully) perform an analysis on his digital device(s). Three considerations are peculiar to network investigation media analysis. The first is the importance of establishing an accurate time line for the actions performed on the digital device so that it can be correlated with your overall network trace information and time line to tie the device to the actions indicated by network evidence previously collected. The second is to identify who was using the digital device(s) at the time of the attacks. This can be accomplished by reviewing connection logs, correlating usernames and passwords to individuals, and asking the individual who owns the computer whether he or she did it. Although cybercrime is steeped in technology, never forget that it's people who perform the actions and a well-done interview can provide the critical information to finalize the investigation. Because compromising digital devices to obtain access or steal information is a central component to cybercrime, malware will almost always be present. Malware is commonly protected by anti-forensic techniques and analyzing it is a specialized skill that is usually performed by dedicated IT security researchers. Malware analysis is extremely important to the development of identification and neutralization techniques, so digital investigators will need to concentrate on identifying, isolating, and extracting it for further analysis.

As with media analysis and litigation support, creating a report that communicates the intricacies of the investigation is a significant challenge. The report will need to explain how the initial attack was conducted, trace the path the attacker took to the victim, correlate dates and time with actions on the victim's, attacker's, and intermediate devices (use a diagram), and interpret the evidence collected from the attacker that can be understood by a layman while providing enough technical detail to support the conclusions drawn from the evidence.

Summary

Digital forensics can be described as a discipline that is both art and science. The science aspect is a combination of the research into hardware and software that expands the knowledge of the digital forensic community and the development and use of the specialized hardware and software tools that are used to collect and analyze digital evidence. Many excellent references contribute to this aspect of digital forensics. The art of digital forensics is exemplified by the experience and intuition of the digital investigator, which this chapter has explored by examining the digital forensic principles, environments, and methodologies, with a focus on conducting and managing investigations.

Solutions Fast Track

Digital Forensic Principles

- ☑ Although it has little to do with the technical side of digital forensics, safety is one principle that cannot be waived.
- ☑ Whether you are performing digital forensics for a civil lawsuit, a criminal investigation, or an administrative inquiry, the authenticity and integrity of the evidence you examined will be of critical importance. Your first step is to establish a chain of custody policy for your organization.
- ☑ You should minimize interaction with original evidence.
- ☑ You should use proven tools and know how they work.
- ☑ Make sure you are conducting objective analysis and reporting.

Digital Environments

- ☑ Many digital forensic investigation will at some point interact with the corporate world.
- ☑ Digital investigators need to avoid the misconception that the government information infrastructure is standardized.
- ☑ In an academic environment, investigators can expect to encounter somewhat chaotic information infrastructures and a lack of centralized management.
- ☑ The home environment encompasses everything from grandma's computer to a small-business network. These environments will be greatly determined by the abilities and interests of the owner.
- ☑ Investigations in the Internet environment will be challenging and rely heavily on the ability of the investigator to obtain either cooperation or legal tools to collect evidence.

Digital Forensic Methodologies

- ☑ In the context of digital forensics, litigation support entails the identification, collection, organization, and presentation of immense amounts of digital data.
- ☑ The overwhelming majority of all investigations (civil, criminal, or administrative) involves digital evidence with digital investigators performing media analysis in cases ranging from employee misconduct to murder investigations.
- ☑ Network investigations are focused on collecting and analyzing digital evidence regarding the use of networked digital devices in commission of crimes and unauthorized activities.

Frequently Asked Questions

Q: Should I obtain digital forensic certification?

A: This dates me, but I can recall a time when there weren't any digital forensic certifications. Now certifications have exploded, vendors (Guidance Software, AccessData, etc.), associations (SANS, IACIS, ISFCE, etc.), and the government (DoD, FBI, etc.) all offer them now. Except for the DoD, a certification is generally not required to perform digital forensics. The courts currently do not require a digital forensic certification before presenting testimony, but you may find getting qualified as an expert witness easier if you have a certification. That being said, I am well aware that a certification only confirms that you successfully demonstrated your knowledge by passing an examination at one point in time. Although possession of certification does not convey mastery of the profession, it does establish a baseline of competence. I know some incredibly talented digital forensic analysts who do not have a single certification, but these people have been working in the field for more than 10 years and have exemplary professional reputations. If you are just starting out in the field, I recommend that you put in the effort to earn a certification; at a minimum, the process of studying for and taking the exam will increase your knowledge of digital forensics and demonstrate that you can perform under some moderate pressure, and in this day and age, having a few initials after your name can't hurt.

Q: What's better, open source or proprietary tools?

A: An accurate and reliable tool that you are efficient with is the best tool regardless of whether the source code is available. That being said, if I have the opportunity to choose from multiple tools that are roughly equivalent in accuracy and reliability, I will select an open source tool. Brian Carrier has published an outstanding paper, "Open Source Digital Forensic Tools: The Legal Argument" (www.digital-evidence.org/papers/), that provides an in-depth discussion of this topic.

Q: How will whole disk encryption affect digital forensics?

A: The recent loss or theft of digital devices containing personally identifiable information has led the corporate world and the government to begin implementing whole disk encryption throughout their mobile device inventories. Digital investigators will need to plan for the presence of whole disk encryption during their precollection planning and collection. Properly implemented, whole disk encryption will completely defeat the standard digital forensic procedure of powering down a digital device, obtaining an image, and then analyzing the image. If you encounter a digital device with whole disk encryption implemented that is powered off, your only options are to obtain the password from the owner through persuasion or legal compulsion, obtain an administrative

password from the whole disk encryption vendor (if one exists and you can persuade or compel the vendor to supply it), or, if the digital device is owned by a business or government agency, obtain a recovery password, once again via persuasion or legal compulsion. If the device is powered on, you will need to conduct a live response to collect evidence and evaluate removing the whole disk encryption. If you do confront a digital device with whole disk encryption, don't fixate on it; start thinking about other places or formats your where your evidence might be. With luck, you may just not need the hard drive.

Q: When imaging a hard drive, is it better to clone the drive or create image files?

A: Prior to the advent of Windows-based forensic analysis suites that could analyze disk image files, hard drives were cloned much more frequently, because even if you imaged a drive by creating image files, you were going to have to restore those files to a sanitized hard drive before you began your analysis. In the Linux/UNIX world, however, image files have always been the standard practice, because you could utilize the loopback driver to mount dd images for analysis. Currently, I am not aware of any digital forensic analysis tools that cannot analyze disk images (at a minimum, dd images), so there is no reason I can think of to clone a hard drive instead of creating image files. Of course, no matter which method you use, you'll still need to obtain and compare hash values.

Q: How is virtualization going to affect digital forensics?

A: The ability to abstract computer hardware with software to create virtual computers which can run different operating systems has become a mainstream technology through the development of vendors (VMWare, Parallels, Microsoft, etc.) and the open source community (Linux Kernel-Based Virtual Machines, Xen, Bochs, etc.). The near-term implications for the conduct of digital forensics are positive, and it is now possible to boot disk images and view/interact with them exactly as the owner would. In addition, virtual machines (and their accompanying virtual networks) make excellent sandboxes for malware analysis. As useful as virtual machines are now, they have the potential to complicate digital forensics. Because a virtual machine will not create operating system artifacts or other evidence on the host operating system, the digital forensic examiner will need to discover whether a virtualization program is installed and then perform a separate digital forensic examination on the virtual machine.

Q: As a non-law enforcement digital investigator, what do I do if I come across child pornography during a digital forensic examination?

A: The law is pretty clear on this. Child pornography is contraband and illegal to possess except under an extremely limited number of circumstances. If you uncover what you

suspect is child pornography, my advice is to stop what you are doing, segregate the digital evidence in question, and *immediately* seek legal advice while informing your client or your client's general counsel. Sources of legal advice for you include your personal attorney, the local prosecutor, or the duty Assistant United States Attorney. Once you've informed your client and obtained legal advice, prepare a report that documents your investigation to this point and create a verified bitstream image of the evidence. Do not conduct your own independent investigation into the suspected child pornography and do not transfer the suspected child pornography to someone else (this may open you to a charge of distributing child pornography). When law enforcement officials arrive to take over the investigation, supply them with your report and either the original evidence or a verified bitstream image (don't forget to fill out the chain of custody documents). If your original investigation is going to continue, I suggest the following protocol: Extract the files you anticipate will be relevant, review them to ensure that they do not contain child pornography, and provide law enforcement with a list of those files when you meet to turn over the original evidence and your report.

FAQ

Q: Is there anything wrong with cross-platform forensic analysis?

A: No, there is nothing wrong with cross-platform analysis. Many tools will perform analyses across operating environments, such as EnCase and the Sleuth Kit. However, be aware that each operating environment has idiosyncrasies that are best analyzed within that particular environment, such as the Windows Registry and Macintosh OS X 10.4 plists.

Q: What effect will the increase in storage capacity have on digital forensics?

A: Moore's Law regarding the doubling of computing power every 18 months is frequently invoked to illustrate the rapid pace of technological growth. I'm not aware of a similarly popular law for storage; although the expansion rate of digital storage available to the average consumer is phenomenal, I'm not that old and I can remember when hard drives were 2 GB, not thumb drives! This increase in storage capacity will have positive and negative effects. On the positive side, the potential for deleted material and operating system artifacts to remain on the digital storage increases because there is so much more storage space (this will be moderated by the optimization algorithms of a particular operating system). On the negative side, digital forensic examiners may have to triage their evidence collection protocols (by digital device or by collecting only files) due to a lack of on-hand digital storage. As we enter the petabyte era in research institutions and data centers, bitstream disk images will probably give way to logical file collection.

Chapter 3

Developing an Enterprise Digital Investigative/ Electronic Discovery Capability

Solutions in this chapter:

- **Identifying Requirements for an Enterprise Digital Investigative/Electronic Discovery Capability**
- **Administrative Considerations for an Enterprise Digital Investigative/Electronic Discovery Capability**
- **Identifying Resources (Hardware/Software/ Facility) for Your Team**

☑ **Summary**

☑ **Frequently Asked Questions**

Introduction

The dramatic expansion of technology and the Internet has transformed how enterprises manage their digital information. In the not-so-distant past, information was handwritten, typewritten, or printed from mainframes. This information, in paper form, was then maintained in vast file cabinets that were difficult to search and access. The advent of computing power to individual employees in an enterprise, whether it is a Fortune 500 corporation, a small accounting firm, or a vast government agency such as the U.S. Department of Defense, has altered the landscape of digital investigations and electronic discovery. The vast majority of information now created exists in digital form on some type of computing system.

The technology (laptops, desktops, personal digital assistants [PDAs], cell phones) that makes an employee so productive is the same technology individuals use to conduct activity that is against company policy or against the law. Employees, whether corporate or government, use technology to "surf" for pornography, "run" an unauthorized business on company time, access accounts to which they don't have authorization, or even steal proprietary corporate information.

Corporate investigators, information security specialists, or law enforcement officers need to be able to investigate these types of allegations/violations by identifying, recovering, analyzing, and reporting on the information that is stored on laptops, desktops, cell phones, PDAs, and other similar devices.

On the legal front, landmark changes to the Federal Rules of Civil Procedure (FRCP) in December 2006 have impacted how digital information is managed in civil litigation. The FRCP formalized the role of digital information in a legal environment and mandates that an enterprise have processes and technologies in place to respond to court orders. The rules have formally identified the role of electronically stored information (ESI) and how it will be handled and presented in a judicial setting.

To properly manage the role of digital information, in either an investigative or a legal setting, a digital investigative and/or electronic discovery capability should be developed/expanded by the enterprise. This capability will provide investigators, information security specialists, attorneys, and law enforcement personnel with the ability to address digital information for either investigative or legal purposes.

Many entities, especially large corporations and government agencies, probably have some digital investigative capability that is utilized to support internal investigations. This capability can be as small as an employee with an additional duty requirement or interest who conducts digital investigations with a stand-alone forensic analysis system and digital forensic software. It can be as large as an established unit with formalized policies/procedures, training standards, enterprise equipment, and hardware for conducting large-scale digital investigative support.

Introduced is the concept that the requirement for an electronic discovery capability can be addressed by the digital investigative team. A digital investigative team will utilize many of

the same policies/processes, equipment, space, and so on that is needed to conduct electronic discovery in an enterprise. Many organizations that currently have an internal digital investigative function, especially in the corporate sector, are using this model and absorbing the electronic discovery mission into their current structure. The appropriate changes are made to address electronic discovery requirements.

This chapter will discuss best practices and guidelines on expanding on a current digital investigative and/or electronic discovery capability or establishing a new capability in the enterprise. You can use this information to formulate a framework to present to senior management for concurrence and approval. At times, it will feel like you are climbing a tall mountain with no end in sight, but when you reach the top, the feeling of having developed and implemented a digital investigative and/or electronic discovery capability in the enterprise will be one of great satisfaction.

Identifying Requirements for an Enterprise Digital Investigative/Electronic Discovery Capability

Developing and implementing an enterprise digital investigative and/or electronic discovery capability can be a costly, time-consuming, and resource-intensive proposition for any enterprise. An organization needs to have a firm understanding of its requirements in this area before embarking on the path to create an internal capability to conduct digital investigations and/or electronic discovery. It is neither inexpensive nor easy to develop and implement this capability in an organization. After careful consideration and deliberation, if the decision is made to move forward on this path, numerous considerations will need to be identified and managed for the successful development and implementation of a program.

The organizational requirement to provide digital investigative support is normally structured so that either the corporate security department or the IT security organization in the enterprise provides this capability. Digital investigative support may be provided to the legal department, human resources department, or ethics department when there is an issue of an employee violating company policy and corporate digital equipment was used, that is, a computer or cell phone was an instrument in the violation. The digital investigative team will need to acquire, analyze, and report on their findings and provide that information to the requesting department for administrative or civil action.

The legal department will be primarily responsible for responding to plaintiff counsel on electronic discovery issues that impact the enterprise. The legal department typically does not have the internal capability to collect, analyze, and report on the data that is being requested in litigation. The legal department will ask the corporate department with the digital investigative capability to provide support on electronic discovery matters (unless the work is being outsourced to a third-party supplier). This takes place because, for the most part, the policies

and procedures utilized for digital investigations will be similar for electronic discovery efforts. It would not make much sense for the legal department to build this capability if it was already available somewhere within the enterprise.

In the government arena, a law enforcement agency will typically be responsible for conducting digital investigations to support or assist in criminal investigations, but not for civil litigation cases. There is a separation in this function due to the sensitive nature of digital criminal investigations. Electronic discovery efforts in government agencies are typically addressed by the IT department which is coordinated by the legal department.

If your enterprise does not have the internal capability to conduct digital investigations and/or electronic discovery, you should decide whether it would be appropriate to develop the capacity to conduct both digital investigations and electronic discovery at the same time. If developed during the same time frame, you will gain "economies of scale" by developing mutual processes, acquiring equipment (hardware/software) that can be used for both digital investigations and electronic discovery, acquiring a facility, training personnel, and so on.

Planning will be a key consideration in developing this capability. If not planned correctly, the development and implementation could have significant negative implications for an enterprise. The digital evidence in a civil or criminal investigation could be barred from use in judicial or arbitration proceedings because of questions of authenticity of the evidence, chain of custody, and so on.

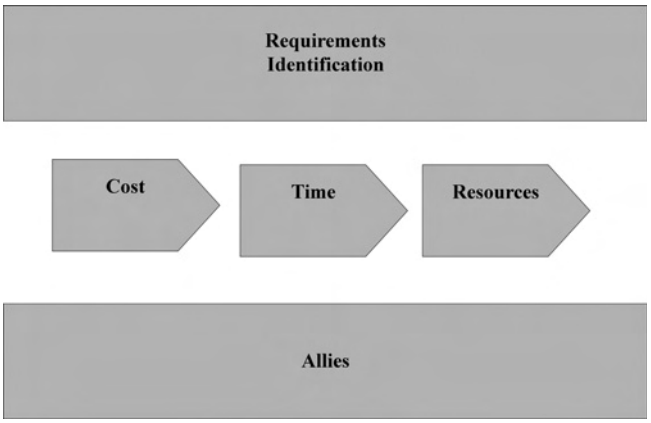
The metadata (data about data) of digital information acquired and analyzed in electronic discovery for civil litigation may be found to be unreliable because timestamps and date stamps were unknowingly altered when the information was improperly collected by an IT employee without proper training or supervision.

Factors that will have a role in the decision to develop the internal capability to conduct digital investigations and/or electronic discovery support will be cost, time, resources, and allies. A business case, which is supported by the requirements identification, needs to be presented to enterprise management that will explain some of the following:

- **What is the problem?** This would explain the problem you are trying to solve for the enterprise (e.g., lack of internal capability requires expensive outsourcing).
- **How you will solve it?** This would explain how an internal digital investigative and/or electronic discovery will solve the problem (e.g., the team can quickly address digital investigative and/or electronic discovery requests).
- **How much it will cost?** This would answer the question of how much more economical it will be to provide the service internally rather than to outsource.
- **What will be the benefits?** This would explain how the internal team will be able to conduct the collection, analysis, and reporting faster, cheaper, and more efficiently, thus minimizing risk to the enterprise.

Identifying requirements will necessitate that several items be considered when putting together a business case for executive management in the enterprise (see Figure 3.1). Items to be considered include costs, time, resources, and allies.

Figure 3.1 Factors to Consider for Developing a Digital Investigative and/or Electronic Discovery Capability in an Enterprise



Costs

Costs need to take into account factors such as time to develop and implement policies and procedures, identification and acquisition of equipment (software/hardware), building a facility, personnel salary, training and certification of personnel, and numerous other items that need to be considered when developing this capability for the enterprise. Identifying whether this will be a small facility with one digital investigator and two or three computers or a large facility with 20+ digital investigators and global operations requires an analysis to determine the type and number of digital investigations and/or electronic discovery requests the facility will manage.

Initiate discussion with your corporate security department (or whichever department has the responsibility for this area) when attempting to identify the number and types of investigations that are conducted that have a digital investigative requirement. Determine how these cases are handled. In some instances, you will find that some corporate security departments do not even know what to do with these types of investigations and thus do not manage them properly. If the department outsources the work, establish the typical cost for collection, analysis, and reporting on digital investigations. This information will be a key consideration in determining the enterprise's expense for this type of work.

The legal department will be the primary focal point for discussion of electronic discovery costs. They will hopefully know the number and types of litigation cases the enterprise is managing. If the work is currently outsourced to a third-party supplier, engage your legal and procurement departments to determine the expense to the enterprise.

The costs can be framed into a return on investment (ROI) to determine the organization's true costs. The ROI will need to take into account the total costs of labor, material, and so on versus the price that would be paid for a third party to provide this service to the enterprise.

Management will need to determine whether the facility will handle only internal digital investigations and not electronic discovery requests. It is less expensive to develop and implement a digital investigative capability without having to conduct electronic discovery for an enterprise because the team will require specialized enterprise-capable forensic and electronic discovery software and hardware, which tend to be very expensive.

A government agency will have different requirements than a corporation. Government organizations that are responsible for law enforcement will traditionally manage only digital investigations that address violations of criminal statutes. As crime continues to evolve in the digital realm, more of the smaller law enforcement agencies are determining that they need to be able to manage digital investigations in their jurisdiction. Electronic discovery support requests in government agencies are typically managed by other organizations in the agency, such as the IT or legal department.

It some instances, it may be more cost-effective or practical for the enterprise requirements for digital investigative and/or electronic discovery support to be met by outsourcing the work to a third-party consultant. A third-party consultant may be able to provide trained and certified individuals along with a facility in which digital evidence/electronic discovery is collected and processed.

Due diligence will be required when outsourcing to ensure that the third-party consultant has personnel that are properly trained and experienced. Experience can make a tremendous difference between an individual that has taken a few seminars and knows only how to use a forensic application and an individual who has been certified, has practical experience in corporate, law enforcement, or military investigative agencies, and has conducted large numbers of digital investigations.

Time

Time is a consideration that needs to be taken into account when determining the viability of an internal digital investigative and/or electronic discovery program. It may take more than a year to plan, develop, and implement a program due to issues with policy/process implementation, funding, training, and resource constraints.

If your organization cannot make the time commitment to develop and implement the program correctly, it may make more sense to outsource to a company that can provide this type of support to your enterprise. This is important because if the time commitment does not allow your organization to develop and implement this capability effectively, you should not take the path to failure by not implementing correctly.

Failure can come in the form of having incorrect policies/processes in place, personnel that are trained incorrectly, or equipment that is inappropriate for collecting or processing

digital evidence. The time commitment must come from senior management who understand that time will be necessary to develop and implement this capability.

Resources

Resources are crucial for developing and implementing a digital investigative and/or electronic discovery program. If senior executives do not understand and "buy into" the concept of such a program, it will not be possible to get the resources that will be vital to the program's success. Resources will be required to manage the process of establishing the capability, to man and staff the organization, to define cost models, to acquire software/hardware, to build of a facility, and so on. The development and implementation are not easy and will require commitment and support from many departments within the enterprise.

It makes sense at this stage to preliminarily engage your organization's project management department or similar group and advise them of the support that will be required for this task if management endorses the program. A project manager could tentatively be assigned that can assist in early resource and requirements identification.

A good project manager can help plan the development and implementation cycle in a methodical manner. The project manager can establish the time frame and cost requirements and ensure that timelines will be met by the responsible parties. Ultimately, the project manager will keep you on track and help make this program a success. Without having a firm understanding of the required resources, it will be difficult to establish the time and cost for developing and implementing the program.

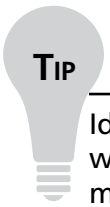
Allies

Lastly, allies must be identified that will promote the requirements to develop and implement a digital investigative and/or electronic discovery capability in the enterprise. Crucial allies will be attorneys from your legal department, and personnel from the human resources department, the corporate security department, and especially the finance department.

Each of these units will have a stake in your success: the legal department because of the services you will provide in the electronic discovery arena, the human resources and corporate security departments because of the support you will provide in assisting their investigations, and the finance department because of the cost savings you will bring to the enterprise by providing this support. In the end, it boils down to costs for the finance department, and if they also agree in the cost-saving measures, they will be a strong ally in identifying funding.

To gain executive support for an internal digital investigative/electronic discovery capability, the premise can be made that an internal capability will be more cost-effective and

reliable than to outsource to a third-party vendor. There are, of course, exceptions to this; in some circumstances, a neutral third party will be required and may be assigned by a judge or special master.



TIP

Identify a senior executive early in the requirements identification stage that will assist in navigating the internal enterprise politics. A senior executive can make it easier to promote the idea of building an internal digital investigative and/or electronic discovery capability. This individual can also provide guidance regarding how to navigate potential roadblocks that may delay or hinder moving forward with this initiative.

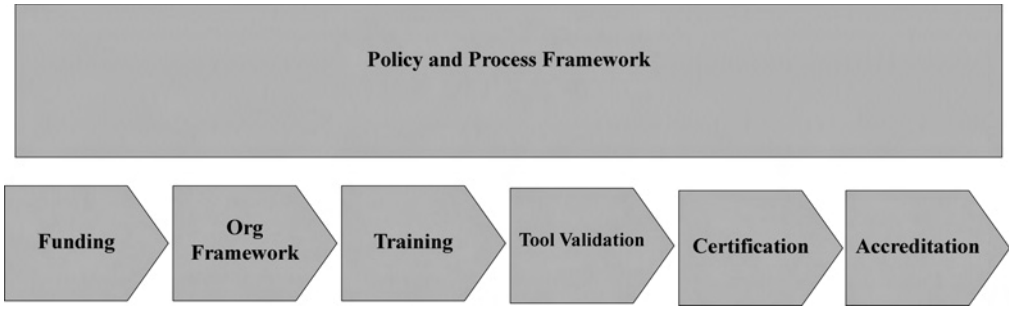
If you have presented a compelling narrative in the business case for developing this capability internal to the enterprise, you will hopefully receive the authority to move forward on implementation and development of the program. Once management has decided to sponsor the development of this internal capability, requirements for the program must be formally defined by an implementation team. Because preliminary work has been conducted up to this stage in the area of costs, time, and resources, it is a matter of adding the previously conducted analysis to the overall program plan.

The implementation team should be composed of individuals that have a firm knowledge and understanding of the role of digital evidence and electronic discovery. A good project manager should also be included in this team. The implementation team will be responsible for defining the equipment (software/hardware) that will be used, the personnel that will be recruited and hired to conduct the work, requirements for the facility to include security, air conditioning/cooling, power, and so on.

Administrative Considerations for an Enterprise Digital Investigative/Electronic Discovery Capability

The administrative considerations to implementing a digital investigative and/or electronic discovery capability require careful thought and deliberation. These considerations will be similar regardless of whether the facility is small or large. The framework and structure need to be in place as soon as possible before any collection or analysis of digital information takes place in the enterprise. The framework should incorporate the items shown in Figure 3.2.

Figure 3.2 The Framework for a Digital Investigative and/or Electronic Discovery Capability



Policy and Standard Operating Procedures

The most critical component to be created early in the development stage of the facility is the policy and process framework.

The policy will provide the enterprise with the necessary authority for managing the digital investigative and/or electronic discovery process in the organization. The policy should clearly identify the department responsible for managing digital investigations and/or electronic discovery in the enterprise. The policy will provide information on the five Ws; who, what, where, when, and why, as shown in Table 3.1.

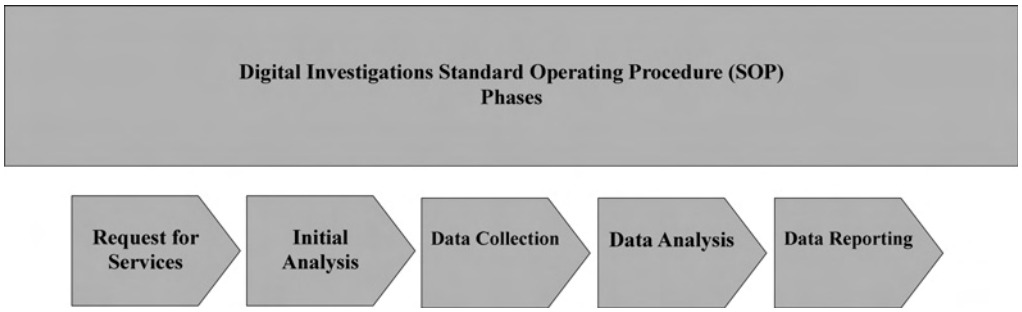
Table 3.1 The Five Ws of a Digital Investigation

Who	What	Where	When	Why
Identify who is responsible for managing the digital investigation/ EDiscovery capability (e.g., corporate security or information security personnel).	Identify what the department is responsible for managing (e.g., conducts digital investigations and electronic discovery).	Identify where the department will provide support (e.g., for the following authorized unites, human resources, ethics, legal, etc.).	Identify when the department will provide digital investigation/electronic discovery support (e.g., assistance will be provided when it is believed that employees are in violation of company policy).	Identify why there is a need for a digital inves tigation/ EDiscovery capability (e.g., provide assistance to authorized departments in order to protect the enterprise, identify unauthorized employee behavior, etc.).

Equally important is the process framework for the team. The process will provide clear details regarding how digital investigations and/or electronic discovery must be conducted in the enterprise in order to have a defensible methodology in place. The process framework details the standard operational procedures that must be observed when conducting digital investigations and/or electronic discovery for the enterprise.

Digital investigations are conducted when there is a violation of company policy or unauthorized activity in a corporate environment. In a law enforcement environment it is driven by a violation of the law, that is, committing cyberfraud, exchanging child pornography, software piracy, and so on, as shown in Figure 3.3.

Figure 3.3 The Phases of the Standard Operating Procedures of a Digital Investigation



Digital investigations Standard Operatiing Procedures (SOPs) should, at a minimum, provide the following phases for the team:

- **Initiation/request for services** Describes the manner in which the enterprise's human resources department, corporate security department, legal department, and so on will request digital investigative and/or electronic discovery assistance and support. The request could be as simple as a phone call or an e-mail sent to a group mailbox. It can be as elaborate as a Web-based request form that is tied to a workflow process that is available on the enterprise Web site. The request must be tracked and should be used to build metrics on how many investigations are initiated, by which departments, the types of cases, and how quickly the team addresses them.
- **Initial analysis phase** Describes the initial information gathering that needs to take place to identify the number of digital investigators, type of software/hardware, and so on that will be required before any data collection efforts can occur in the enterprise. This phase includes the following stages:

Documentation Memorize your actions and activity so that if the case is presented to an attorney, jury, and so forth, you can explain the actions that occurred in the investigation. An individual that sues your enterprise on an allegation of unlawful termination, for example, may request copies of the information that was used to make the decision to terminate. This information may be discoverable in legal proceedings. It may be years before the case is presented to an arbitrator or a court for their opinion on the matter. The team will need to be able to provide the documentation that describes your department's actions. Ensure that your documentation is clear and concise and that it describes only the facts of the investigation. Do not speculate or theorize on the documentation; again, this information may be discoverable and you may have to explain why you speculated based on comments in your notes.

Planning Engage with the department (human resources, corporate security, legal, etc.) requesting support to ensure that all parties agree with the services being requested. During these sessions, an understanding is made of the type of information the digital investigative team will acquire. Keywords, file types, date and time, and so on will be agreed to during these meetings. This is an important time to narrow down the scope of the digital investigation so that it is manageable. The amount of data that is contained in enterprise hard drives is just too large for a digital investigator to "find evidence." The examiner will need to conduct a focused search for the information that is being requested. This can be accomplished only by working with the requesting department and getting the information you need to be both effective and efficient. Make sure a planning meeting is conducted when a request for digital investigative services is made in the enterprise.

Identification Categorize the types of computer systems (laptop, desktop, server, etc.) that will be the target of the collection effort. If this is a large collection effort, it will probably require collaboration with the IT department for network topology and data maps. This will also assist in establishing an initial baseline on storage capacity that the team will require.

- **Data collection phase** Describes how data will be acquired from different types of digital information sources. Data collection is the process of acquiring digital information in a manner in which the integrity and authenticity of data that is collected is maintained. This phase includes the following stages:

  Documentation Similar to the initial analysis phase, memorize your actions so that if the case is presented to an attorney, jury, and so on, you can describe the activity that occurred in the investigation. This is also the phase where chain of custody begins. Chain of custody is crucial to maintain and provides the digital

examiner with the information necessary to explain when a system was acquired, by whom, and where the data was stored in the enterprise. If chain of custody is broken or unclear, it may become an issue in getting the data that you collected to be allowed in legal proceedings.

CMOS date/time This is an important step in a digital investigation where date and time creation of files is in question. The CMOS will need to be queried before any imaging takes place to determine the date and time on the system to be analyzed. Document this information.

Imaging protocol It is imperative that you acquire a forensic image (bit-by-bit stream) of any data that the team will collect. Two kinds of images can be made: a physical (bit-by-bit) image of the entire physical device, or a logical image which comprises the active directories and files that are available to the operating system. A physical image is the standard for data collection in a digital investigation. A forensic or bitstream image captures all the data contained on the device that is being collected, to include deleted or remnants of information.

An enterprise will have a large and diverse number of computing devices on the network. The team should acquire knowledge and experience in imaging these devices before an actual support request comes in to capture the data on these systems. Examples of the different types of data devices are:

Large data set collection Redundant Arrays of Inexpensive Disks (RAIDs) in servers, storage area networks (SANs), and so on

Desktop/laptop collection Personal computing devices

Mobile device collection PDAs, cell phones, USB drives, and so on

Coordinate with the IT department to receive assistance on how these devices work and ask them to provide information on any new IT systems that are brought into the enterprise environment. A good working relationship with the IT department is critical.

Hashing This validates that the digital investigator has not altered the integrity of the acquired data. A digital hash conducts a mathematical algorithm of a device or files and provides a digital fingerprint. The digital fingerprint is used to authenticate that the digital investigator has not tampered with or altered the captured data. This digital fingerprint information is always maintained with the case file and provided if there are questions about the image's authenticity.

- **Data analysis phase** Describes how data will be processed for information that will be of pertinent value for the requesting department. A methodology or checklist should be established that will guide the digital investigators on how to conduct data analysis. The guidelines should not be so detailed that the investigator

is not provided the leeway to use his background and experience to identify information that may be of value in the investigation. It should serve as a guideline and provide a process framework that should be followed along with other techniques that the investigator may feel is useful in the data analysis phase. At a minimum, the guideline or checklist should require the investigator to conduct the following tasks:

Hash analysis search Similar to the hashing that takes place in the data collection phase, hashing can be utilized to search for and identify responsive files. Responsive files are files that are of pertinent value to the requesting department. Hash function software such as an SHA-1, MD5, or CRC application can be utilized to conduct a mathematical analysis of a data storage device (hard drive, USB, etc.) or file and the output is a digital fingerprint of the data that was hashed. The digital examiner would compare the hash fingerprint against the hash values on the data that he is examining.

Hashing will also enable the digital investigator to exclude large numbers of files that have no value to the case. It is also highly recommended that known system and application files be excluded from review. The National Institute of Standards and Technology (NIST) makes available hash sets of known files for hundreds of different applications and operating systems. The NIST Web site, www.nsrl.nist.gov/index.html, provides downloadable CDs that contain SHA, MD5, and CRC values for more than 42 million files. You can easily exclude from analysis files whose hashes are known so that the team can concentrate only on pertinent information in the case.

Enterprise forensic software applications such as AccessData's Forensic ToolKit (FTK), Guidance Software's EnCase, and Technology Pathways' ProDiscover can hash files that have been imaged for analysis. The process utilized to conduct a hash must be repeatable to ensure that the results of the device or file that is hashed can be reproduced for authenticity of digital evidence purposes.

Keyword search A search can be conducted by utilizing keywords that have been provided by the requesting department or that have been identified through investigation. A keyword search will narrow the scope of the analysis to only those files that are responsive to the search. A good keyword list will save a lot of time and effort and allow the digital investigator to concentrate on the information that is of value.

The keywords and/or keyword list should be acquired in consultation with the department requesting the support. They will be most familiar with the case and what they need for their investigation. It is important to meet with the requestor early in the initiation of the investigation to acquire this information.

File signature search A search can be conducted by utilizing file signature analysis. A file signature can be used to narrow down the types of files that are of interest in the investigation. For example, if the human resources or legal department is asking for Microsoft PowerPoint files, it will be a matter of isolating only those files that are of interest to the requesting department and not reviewing Excel files.

- **Data reporting phase** Describes how the data acquired during the collection and reporting phases will be presented. A report is the final product that will be provided to the department that requested the digital investigative support from your team. It is important that the report only describe facts that were observed, analyzed, and thus documented. As a digital investigator, it is crucial that the information reported speak only to the facts of the case. The report should be a neutral and detached observation on the information that was observed during the analysis. If you put information subjectivity into the report, you are no longer being an objective reporter. The most important role that a digital investigator and a digital investigative and/or electronic discovery facility can bring to the enterprise is neutrality in capture, analysis, and reporting.

  The report must contain at least the following information:

  Executive summary Describes the facts of the case and the information that was observed, and is pertinent to the investigation.

  Analysis summary Describes in detail the collection and analysis that the digital investigator conducted. This is a very thorough and complete description of the steps the investigator took to prove or disprove the allegations that he is investigating for the requesting department.

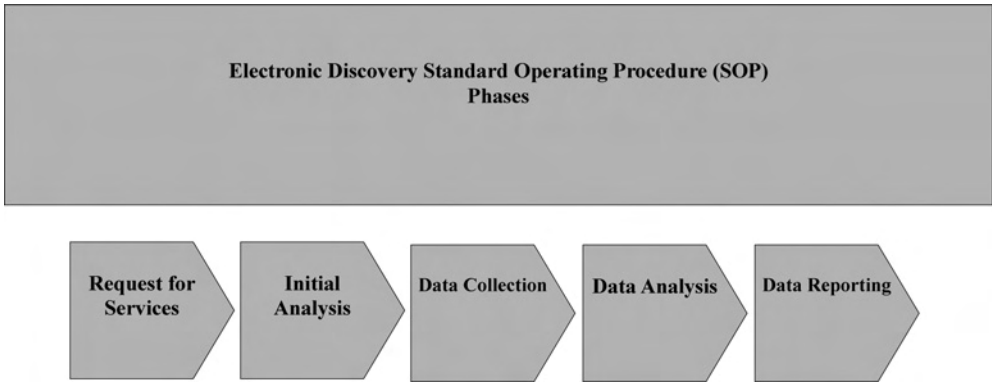
  Final summary Describes the analysis and conclusion.

Electronic discovery is driven by external legal or regulatory action directed at an enterprise. The corporation/agency will be responding to requests from attorneys or regulators to provide information. Recent changes to the FRCP have added more rigors to the protocol used when discovering electronic information that will be presented in federal court. The purpose of the changes to the rules is to ensure that litigants requesting or responding to electronic discovery requests are operating from the same set of standards.

For enterprises to respond to requests for ESI, which can be defined as e-mails, Web pages, laptop/desktop/server hard drives, and/or CD/DVD database data that is stored by or on behalf of the corporation/agency, processes must be in place to identify and deliver this information to a requesting party. The procedures must be clearly defined and defensible so that the digital investigations and/or electronic discovery team will be in a position to attest to the discovery that the organization conducted.

The team can use the SOPs shown in Figure 3.4.

Figure 3.4 The Standard Operating Procedures for the Phases of Electronic Discovery



Electronic discovery SOPs will typically follow the same process as the digital investigative support process, but with some differences. The distinction will be primarily in the data collection phase. Data collection for electronic discovery will be more complex due to the legal department's close involvement in the effort. The electronic/digital systems that will need to be searched for responsive data will also tend to be much larger than those in a digital investigation.

The electronic discovery SOP should, at a minimum, provide the following phases for the digital investigative and/or electronic discovery team:

- **Initiation/request for services** Describes the manner in which the organization's legal department will request electronic discovery assistance and support. The request can be as simple as a phone call or an e-mail sent to a group mailbox. It can be as elaborate as a Web-based request form that is tied to a workflow process that is available on the enterprise Web site. This form should have a tracking mechanism that can be tied to metrics that are used to establish efficiency of the team in responding to the legal department's electronic discovery requests.
- **Initial analysis phase** Describes the initial information gathering that needs to take place to help identify the number of digital examiners, types of software/hardware, and so on that will be required before any data collection efforts can occur in the enterprise. This phase includes the following stages:

  Documentation Memorize your activity so that if the case is presented to an attorney, jury, and so on, you can explain the actions that occurred in the electronic discovery effort. An individual that sues your enterprise on an allegation of unlawful termination, for example, may request copies of the information that was used to make the decision to terminate. This information may be discoverable in legal proceedings. It may be years before the case is presented to

an arbitrator or a court for their opinion on the matter. The team will need to be able to provide the documentation that describes the actions that your department took. Ensure that your documentation is clear and concise and that it describes only the facts of the investigation. Do not speculate or theorize on the documentation; again, this information may be discoverable and you may have to explain why you speculated based on comments in your notes.

Planning Engage with the legal department to ensure that all parties agree on what the team is requesting. During these sessions, an understanding is made of the type of information that the team will acquire. Keywords, file types, date and time, and so on will be agreed to during the meeting. This is an important time to narrow down the scope of the electronic discovery support. The digital investigator will need to conduct a focused search for the information that is being requested. This can be accomplished only by working with the requesting department and getting the information you need to be effective. At this point, the team may also engage the IT department for support on enterprises systems that they may have specialized knowledge on recovering data.

Identification Categorize the types of computer systems (laptop, desktop, server, etc.) that will be the target of the collection effort. If this is a large collection effort, it will probably require collaboration with the IT department for network topology and data maps.

- **Data collection phase** Describes how data will be acquired from different types of electronically stored information sources. This phase includes the following stages:

  Documentation Similar to the initial analysis phase, memorize your activity so that if the case is ultimately presented to an attorney, jury, and so on, you can explain the actions that occurred in the electronic discovery effort. This is also the phase where chain of custody begins. Chain of custody is crucial to maintain and provides the digital examiner with the information necessary to explain when data was acquired, by whom, and where the data was stored in the enterprise. If chain of custody is broken or unclear, it may be an issue in getting the data that you collected to be allowed in legal proceedings.

  CMOS date/time The CMOS may need to be queried before any imaging takes place to determine the date and time on the system to be analyzed. This step may not be required but should be discussed with the legal department to make sure they understand how date/timestamps may impact the electronic discovery effort.

  Imaging protocol It may not be necessary to acquire a forensic image (bit-by-bit stream) during an electronic discovery collection effort. Two kinds of

images can be made: a physical (bit-by-bit) image of the entire physical device, or a logical image which comprises the active directories and files that are available to the operating system. There is currently no established legal ruling or precedent on how data will need to be captured or imaged for electronic discovery.

What is not in dispute is that the metadata (data about data) of files must not be modified, changed, or altered. Thus, it is permissible to capture the responsive data by searching in the logical files in the systems that have been identified as potentially containing data of interest. The critical consideration is that no data elements are changed when the data is captured for analysis. Remember to work closely with your legal department for opinions on the collection efforts that the team will conduct.

An enterprise will have a large and diverse number of computing devices on the network. It is recommended that the digital investigative team acquire knowledge and experience in imaging these devices before an actual electronic discovery support request comes in to capture the data on these systems.

This is especially important in electronic discovery efforts because of the time constraints that are imposed by the legal department and their requirement to respond to opposing counsel.

Examples of the different types of data devices are:

Large data set collection RAIDs in servers, SANs, and so on

Desktop/laptop collection Personal computing devices

Mobile device collection PDAs, cell phones, USB drives, and so on

Coordinate with the IT department to receive assistance on how these devices work and ask them to provide information on any new IT systems that are brought into the enterprise environment. Data maps are important and can provide a topology of how electronically stored information is maintained in the enterprise.

Hashing This validates that the digital investigator has not altered the integrity of the acquired data. A digital hash conducts a mathematical algorithm of a device or files and provides a digital fingerprint. The digital fingerprint is used to authenticate that the team has not tampered with or altered the captured data. This digital fingerprint information is always maintained with the case file and provided if there are questions about the authenticity of the acquired data.

- **Data analysis phase** Describes how data will be processed for information that will be of pertinent value for the requesting department. A methodology or checklist should be established that will guide the team on how to conduct data analysis. The guidelines should not be so detailed that the digital investigator is not provided the leeway to use his background and experience to identify information that may be

of value in the electronic discovery effort. It should serve as a guideline and provides a protocol that should be followed along with other techniques that the digital investigator may feel are useful in the data analysis phase. At a minimum, the guideline or checklist should require the digital investigator to conduct the following tasks:

Hash analysis search Similar to the hashing that takes place in the data collection phase, hashing can be utilized to search for and identify responsive files. Responsive files are those files that are of pertinent value to the requesting department. Hash function software such as an SHA-1, MD5, or CRC application can be utilized to conduct a mathematical analysis of a data storage device (hard drive, USB, etc.) or file and the output is a digital fingerprint of the data that was hashed. The digital examiner can compare the hash fingerprint against the hash values on the data being examined to validate that the data has not been altered.

Hashing will also provide the investigator with a method to exclude large numbers of files that have no value to the electronic discovery effort. It is also highly recommended that known system and application files be excluded from review by the digital investigation and/or electronic discovery team. NIST makes available hash sets of known files for hundreds of different applications and operating systems. The NIST Web site, www.nsrl.nist.gov/index.html, provides downloadable CDs that contain SHA, MD5, and CRC values for more than 42 million files. You can easily exclude from your analysis files whose hashes are known so that the team can concentrate only on pertinent information in the case.

Enterprise forensic software applications such as AccessData's Forensic ToolKit (FTK), Guidance Software's EnCase, and Technology Pathways' ProDiscover can hash files that are imaged for analysis. The process utilized to conduct a hash must be repeatable to ensure that the results of a device or file that is hashed can be reproduced.

Keyword search A search can be conducted by utilizing keywords that the requesting department has provided or has identified through investigation. A keyword search will narrow the scope of the analysis to only those files that are responsive to the search. A good keyword list will save a tremendous amount of time and effort and allow the digital investigator to concentrate on the information that is of value to the electronic discovery effort.

The keywords and/or keyword list should be acquired in consultation with the legal department which is requesting the support. They will be most familiar with the case and what they need for their discovery. It is important to meet with the legal department early in the investigation to acquire this information.

File signature search A search can be conducted by utilizing file signature analysis. File signatures can be used to narrow down the types of files that are of interest in the electronic discovery effort. If the legal department is asking for Microsoft PowerPoint files, for example, it will be a matter of isolating only those files that are of interest to the requesting department and not reviewing Excel files or other nonpertinent files.

- **Data reporting phase** Describes how the data acquired during the collection and reporting phases will be presented. A report is the final product that will be provided to the legal department that requested the electronic discovery support from your team. It is important that the report only describe facts that were observed, analyzed, and documented. As a digital investigator, it is crucial that the information reported speak only to the facts of the case. The report should be a neutral and detached observation of the information that was observed during the analysis. If you put subjectivity into the report, you are no longer being objective in your collection and analysis. The most important role that a digital investigator and a digital investigative and/or electronic discovery facility can bring to the case is neutrality in capture, analysis, and reporting.

  The report must contain at least the following information:

  Executive summary Describes the facts of the case and the information that was observed, and is pertinent to the electronic discovery effort.

  Analysis summary Describes in detail the collection and analysis that the digital investigator conducted. This is a very thorough and complete description of the steps the investigator took to identify, capture, preserve, and report on the findings of the electronic discovery effort that was conducted in the enterprise.

  Final summary This comprises the analysis and conclusion.

Funding

Funding the initial costs of developing a digital investigative and/or electronic discovery capability is only the first step in the development and implementation cycle. A sustainable budget model should be developed that projects the requirements for equipment, personnel training, and software in three-year increments. It is important to identify and consider additional equipment, certifications, and so on that will be required during the time frame in consideration by the enterprise.

Technology is changing at a very rapid rate and the software and hardware that have been purchased for collection and analysis of digital data may be obsolete during the current budget cycle. It is thus crucial to make educated assumptions on forensic software and hardware that may be required and how to fund it with available or supplementary budgets.

A good source for an enhanced budget may be the legal department, especially if large electronic discovery support efforts are being provided by the team.

Several sustainment budget models should be considered. They include:

- **Department-funded model** The digital investigative and/or electronic discovery team is funded by the enterprise to provide internal support services and identifies the budget that is required to fund this type of work. This is the most efficient budget model for an organization because you control the budget and do not have to charge back to the requesting department on support that is requested.

  An essential requirement will be to track time, personnel, and equipment resources that are utilized for digital investigations and electronic discovery efforts. These costs will be used to track the effort that your team provides to departments in the enterprise. As the workload grows, and it will, you can use these metrics to request enhancements to the budget for the team.

- **Bill-back model** The team bills the requesting department for support services that are provided. The team will need to develop a cost code and price model for the services that it will provide in the enterprise.

  The price model can be as simple as an hourly fee that will be charged to all departments in the enterprise for digital investigative and/or electronic discovery services. It can be as complex as individual fees that are charged for the different tasks that are provided for the requesting departments. There could be a fee for collection that is different from a fee for processing, for example.

  This model is harder to implement in the enterprise because the department that you will be providing support to may not know how much of your services they will require. Not knowing what they will need in a given year, it will be hard for them to project a budget to which your team can bill back.

Organizational Framework

The organizational framework will define how the team will be structured in the enterprise. If the team will be part of the corporate security organization, the role must be defined for how this unit will interact with others in the enterprise.

The structure will also establish how cases will be brought into a workflow process for the enterprise. A workflow process is necessary for efficiency and will ensure that requests for digital investigative and electronic discovery support are conducted in a timely manner. The workflow will ensure that assignments are completed as soon as possible and will follow an established process.

An information or case management tool must also be used for managing the status of the cases the team will address.

Training

Training is required for any individual that will be collecting or processing digital information. Training will provide the core skill set that every member of the team will need before they can be permitted to conduct digital investigations and/or electronic discovery efforts.

Training should provide the following minimum skills:

- Handling digital evidence and chain of custody
- Basic digital forensics and investigations
- Advanced digital forensics and investigations
- Electronic discovery

The team's management must not allow an individual to conduct digital media collection or analysis without having the proper training on the use of forensic/electronic discovery software. It is highly recommended that the SOPs define the training requirements that each individual on the team must meet.

Tool Validation

The purpose of tool validation is to ensure that the software being utilized maintains data integrity. Tool validation is needed before any digital examination takes place in the enterprise.

A secondary and equally important consideration in an enterprise is to have the hardware and software tested to confirm that the introduction of this technology will not have a negative impact to the enterprise network. It is important to engage your IT staff to advise them of your objective and the requirements to install enterprise tools—for example, servers that will process and store large amounts of data, software (agents) that will be placed on enterprise computers (laptops, desktops, servers), and so on.

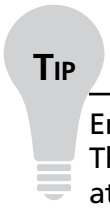
To validate the tools that your department will be using, the following methodology may be beneficial:

- Prepare a USB drive:
  1. Wipe the USB of all data with a wiping utility such as AccessData's WipeDrive utility (the Format utility that Microsoft Windows provides does not wipe the data from the device; it just removes the directory pointer, and thus, you should not use this application to wipe devices).
  2. Format the USB drive.
  3. Validate that there is no data on the USB drive with a disk-editing or forensic application.

4. Create and document a test scenario.
5. Create several folders on the USB drive.
6. Create several Word, Excel, and PowerPoint files that contain data (use whatever information you want) and copy them to the folders on the USB drive.
7. Delete some of the files that are located in the folders in the USB drive.
8. Delete some of the folders that are located in the USB drive.

- Validate with the tool being tested:
  1. Conduct a forensic image of the USB drive with the tool that is being tested.
  2. Analyze the forensic image with the tool being tested.
  3. Identify the folders and files that *were not* deleted and validate that they are still present and not changed.
  4. Identify the folders and files that *were* deleted and validate that the tool being tested recovered these items.
  5. Store the results of the validation effort.
  6. Store the documentation that was created to validate the tool.

The tool validation will reasonably reassure the organization that the software it is using is working as intended.



TIP

Ensure that you validate the forensic tools you will be using in your facility. The examiner does not want to be on the witness stand explaining to an attorney why he knows the tool works. An answer of "I read in the manual that the tool works" is not acceptable.

Certification

Certification of personnel that will acquire, analyze, and report on digital data will be important for the enterprise. It shows that at a minimum, the individual conducting the collection, analysis, and reporting has a met a basic level of understanding of protocols and use of forensic software tools. Several states, including Texas and New York,) now require government officials that conduct digital media analysis for law enforcement purposes to have some type of certification.

General digital forensic certifications, which teach principles and techniques and are not software-specific, can be obtained from several organizations, some of which include the following:

- SANS GIAC Certified Forensic Analyst (GCFA) (www.sans.org)
- IACIS Certified Forensic Computer Examiner (CFCE) (www.cops.org)
- EC-Council Certified Hacker Forensic Investigator (CFCI) (www. eccouncil.org)

Software-specific certifications can be obtained from organizations such as:

- AccessData (www.accessdata.com), which has the AccessData Certified Examiner (ACE)
- Guidance Software (www.guidancesoftware.com), which has the Encase Certified Examiner (EnCE)

Accreditation

Accreditation is a factor that should also be considered in the early stages of developing and implementing the digital investigative and/or electronic discovery team. If policy and procedures are developed early in the process, this will make it a lot easier to go down the path of accreditation in the future. Accreditation in not easy for any enterprise to achieve because of the complexities of process review. It requires a tremendous amount of time and documentation to meet the stringent requirements that the American Society of Crime Laboratory Directors (ASCLD; www.ascld.org) has established.

Identifying Resources (Software/Hardware/Facility) for Your Team

It is important to identify the proper software, hardware, storage, and space that will be used when developing the digital investigative and/or electronic discovery facility. You will need digital forensic software to conduct digital information collection and analysis for the enterprise. Do not rely on any one type of software to conduct your collection and analysis. It is important to have several tools to validate that the information has been properly collected and analyzed. Testing and validation is an important consideration that should be applied often in the enterprise.

It is critical to properly identify the hardware that consists of computer equipment such as servers, desktops, laptops, write blockers, and so on.

Finally, an organization needs a secure facility where the collection, storage, and processing of digital information will take place. This space will need to take into account security, power, and air conducting considerations when it is built. A process must be developed for

controlled entry into this space. Access should be granted only to the individual who will be responsible for evidence control and a backup custodian. Any other individuals that need access to this space must be "signed in" to document that they were in the controlled area.

Software

Digital forensic software is a key component for the team. This specialized software provides the team with the ability to acquire, analyze, and report on information that is of pertinent value. It is important that the software be thoroughly tested to ensure that it meets the established requirements. Digital forensic software can be very expensive for an organization, so you don't want to purchase it and then realize that it will not meet the enterprise's needs.

Several applications are available, each with advantages and disadvantages. It is crucial for the team to test the applications before making a purchase. Some software can be utilized in an enterprise capacity; other software can be used only in a stand-alone environment. It is vital that the team determine what type of application the enterprise needs and then search for the tool that will meet the requirement.

Extensive software testing should take place in a test environment and not in the production network. Identify the most robust scenario where the application would be used and test, test, and test some more. Speak with digital investigative and/or electronic discovery teams in other enterprises and get their thoughts on the applications. Let them tell you what works and doesn't work. Lastly, do not accept what you read on the sales brochure, as this is just marketing material that does not describe every scenario you may encounter in the enterprise.

The following commercial and government forensic software is among that which is available for digital investigative and/or electronic discovery teams:

- Guidance Software EnCase (www.guidancesoftware.com)
- AccessData Forensic ToolKit (FTK) (www.accessdata.com)
- ProDiscover (www.techpathways.com)
- WetStone Technologies Gargoyle Investigator (www.wetstonetech.com)
- ILook (available to law enforcement/military personnel only; www.ilook-forensics.org/)

Hardware and Storage

The team needs hardware, storage, and write blockers to collect and process the data that will be acquired for analysis. If the equipment is not available, it will be impossible for the team to become operational. One of the most important considerations when acquiring equipment is to purchase write blockers to prevent changes to the digital media that the team will collect.

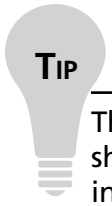
Hardware

The individual examiner should, at a minimum, have the following hardware for collection and digital media analysis:

- 2.2 GHz Intel or AMD processor
- 4GB of RAM
- 1TB SATA hard drive
- 32- or 64-bit Windows XP or Vista
- Video card with minimum 256MB internal RAM
- DVD burner with read/write capability
- FireWire/USB ports
- 21-inch monitor (an examiner will spend a lot of time working on the monitor)
- Speakers

These are only recommendations for hardware. Your best bet is to acquire the fastest computers with the most RAM that you can afford for the facility. Speed and memory are of vital importance for the computer systems that the examiner will use. Soon after you purchase computers for conducting digital media analysis or electronic discovery, these systems will be obsolete.

A procedure should be established so that these computers can be refreshed or refurbished on a regular basis, maybe as often as once a year. Depending on the enterprise's procurement cycle, you may have to purchase the computer systems through your IT organization. If this is the case, engage them early and advise them of your requirements. The department will need to be coached on the type of equipment you need for your team.



TIP

The computer systems utilized to acquire and process digital information should not be the same as those used to check your e-mail and enterprise intranet or Internet. You will be processing data in sensitive investigations or electronic discovery matters and you do not want to alter the data or have its integrity questioned.

Storage

An enterprise will require a large storage capacity to maintain the vast amount of data that the team will collect and analyze. The requirement will be determined by the types of cases the team is working on. Discussions with the IT department are important to determine the team's storage requirements.

Redundancy is of crucial importance in that you don't want a server to fail and not have a manner to recover the data. The storage capacity will also be dictated by how long you will maintain the data that your organization has collected. In electronic discovery efforts, the data may need to be maintained for years. I know of one large enterprise that uses hundreds of terabytes of storage capacity for digital investigations and electronic discovery requirements.

Write Blockers

A write blocker is a mandatory piece of hardware that must always be used when collecting data. Write blockers prevent the examiner from "writing" or "altering" the digital information that is being collected and analyzed. Write blockers are designed and available for IDE, SATA, SCSI, and USB hard drives. They are also available for FireWire and removable media cards such as memory sticks, multimedia cards, and so forth.

Here is a list of some of the companies that sell write blockers:

- Forensic Computers (www.forensic-computers.com)
- WiebeTech (www.wiebetech.com)
- Digital Intelligence (www.digitalintelligence.com)
- Guidance Software (www.guidancesoftware.com)

It will be equally important for the organization to test and validate that the hardware performs as advertised. Testing must be conducted before any collection and processing takes place to validate that the product functions as advertised.

Facility

The facility where you will collect, process, and store evidence and equipment must be planned well in advance of building the space. If at all possible, identify a location that will allow the space to scale as future requirements are identified that will utilize the facility's resources.

It is wise to plan for growth, especially when word gets out in the corporation/agency about your team's ability to provide these important services to the enterprise. It is important to meet and discuss all space issues with the enterprise facility personnel. The facility department will be in a position to physically implement the requirements your organization is requesting. They will be able to assist in identifying a location and building the rooms, and implementing the air conditioning and security features for the space.

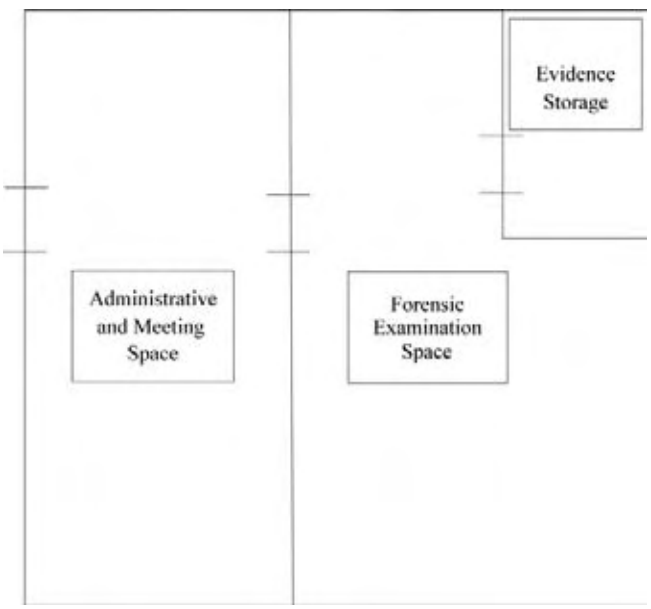
Location

The facility needs to be located in an area where it can be secured from unauthorized access or viewing of the work being conducted in the space.

The overall facility should be segmented into a minimum of four distinct areas (see Figure 3.5):

- **Administrative/meeting area** Managers will utilize this space to conduct briefings on assignments and the status of cases, and to meet with attorneys or other personnel that may have business to conduct in the facility.
- **Examiner work area** This is where the examiner swill spend the majority of their time acquiring digital evidence, conducting analysis, and developing reports on their findings. Each examiner, at a minimum, should have approximately 100 square feet of space for their equipment, books, and so on.
- **Digital evidence storage area** This space must be extremely secure with access to only the evidence custodian and the designated backup custodians. This space will require floor-to-roof walls to ensure that an unauthorized individual cannot climb over the tile on the roof. An alarm system is a must. An information management system must be maintained which identifies where digital evidence is maintained and when it was accessed, by whom, and when it was returned.
- **Network facility area** This space is used to house the servers and network devices that store and transmit information in the facility. There should be two separate and segregated networks for the facility: a network that is used for day-to-day administrative work and a network that is used to collect and process digital evidence.

Figure 3.5 Sample Digital Investigative/Electronic Discovery Facility



Security

The space for the facility must have both physical and logical security access controls to prevent unauthorized personnel from entering the facility. If possible, two-factor physical security controls should be implemented to ensure that only authorized individuals are able to enter the work area where digital information will be collected, processed, and stored. Two-factor authentication can be described as the following:

- **Something you are** Fingerprint scanner, retinal scanner, and so on
- **Something you know** Password, pin, and so on
- **Something you have** Digital token, bingo card, and so on

The two-factor controls can be an identification card along with a keypad entry system or an identification card paired with a biometric system such as a finger or retinal scanner.

At a minimum, any personnel that will have access to the area where collection and processing will take place must sign in and out of the facility. This will ensure that documentation is available that registers any individual allowed access to the secure area.

Ventilation and Air-Conditioning Systems

Ventilation and cooling are crucial to ensure that computer systems that are collecting and processing digital information do not get overheated and shut down or fail. There is nothing worse than to have conducted a large collection or processing requirement that has taken several days/weeks and to have the systems shut down due to lack of circulation and cooling in the facility.

It is important to work with your organization's facilities personnel to identify the appropriate ventilation and cooling requirements. Ensure that any ventilation or air conditioning hardware that will need to be put in place is installed before the facility is cleared as a work space. If there is not enough space or the facility does not have the physical configuration to allow the appropriate ventilation and cooling systems, the organization will need to identify another location that will be more suitable for digital evidence collection and processing.

Electrical and Power Systems

A facility in which several computers/servers are utilized to collect and analyze data will need to have the appropriate power output to sustain the operation of these systems. The hardware utilized in the facility uses a tremendous amount of electricity and management needs to ensure that there is sufficient power to prevent overheating and damage to this expensive equipment.

The electrical and air conditioning requirements must be identified well in advance in the initial planning stage. If not, it is possible that a facility will be built and you will quickly realize that there is not enough electrical output to keep the systems running when heavy acquisition and processing are taking place.

If at all possible, the facility should be on a high floor so that severe rain or flooding cannot damage or destroy equipment and digital evidence.



TIP

Early in the development phase, identify the ventilation, cooling, and power requirements for your facility. Provide information on the types of systems that will be placed in the facility and how long they will be collecting and processing. Most important, plan for future growth.

Summary

An enterprise, whether it is a corporation or a government agency, that has developed and implemented an internal digital investigative and/or electronic discovery capability will be in a better position to address issues of employee misconduct where an individual has used internal computing systems, conduct criminal investigations where a suspect has used a computer, or respond to litigation requests from plaintiff attorneys or judicial authorities. As we described in this chapter, it is neither easy nor inexpensive to build the capability to conduct digital investigations or electronic discovery in the enterprise. But is it a necessity.

It is important that discussions be held and management be made aware of the implications of not having the capability to provide this support. A cost-benefit analysis must be conducted to determine whether the capability is warranted in-house or if the work should be outsourced. Not having an internal capability or process for working with an external third party that provides this type of service is not an option.

The use of technology and the Internet by employees in an enterprise demands that corporations/agencies develop and implement a methodology for addressing these highly complex investigations or electronic discovery requests.

The digital investigative and/or electronic discovery team will be enormously satisfied when this capacity has been successfully executed in the enterprise.

References

Here is a list of documents and URLs where you can acquire additional information on developing and implementing a digital investigative/electronic discovery capability in the enterprise:

- Forensic Laboratories: Handbook for Facility Planning, Design, Construction, and Moving (www.ncjrs.gov/pdffiles/168106.pdf)
- NIST Computer Forensics Tool Testing Project (www.cftt.nist.gov/)
- (NIST National Software Reference Library (www.nsrl.nist.gov/index.html)
- International Association of Computer Investigative Specialists (IACIS) (www.iacis.com/iacisv2/pages/forensicprocedures.php)
- International Journal of Digital Evidence (www.utica.edu/academic/institutes/ecii/ijde/)
- Digital Forensics Research Workshop (www.dfrws.org/)
- American Society of Crime Laboratory Directors (www.ascld.org)

Frequently Asked Questions

Q: A large corporation may process and store hundreds to thousands of terabytes of data. Is it feasible to make a "physical image" of hundreds or thousands of hard drives in an electronic discovery effort?

A: As the amount of data that an enterprise produces continues to increase, the practical implications of this recourse are becoming extremely limited. It is no longer a viable option to forensically image hundreds or thousands of hard drives. When the plaintiff and the defense attorneys confer, they will need to agree on keywords, file types, date ranges, and so on that the defendant's digital investigative/electronic discovery personnel will search. Once the team identifies the responsive files, they are collected and preserved and then provided to the enterprise attorney for review and further action.

Q: How much does it cost to outsource digital investigative or electronic discovery work to an external company?

A: The costs typically vary with individual vendors. It will depend on the time and effort required for the effort. In electronic discovery cases, due to the large amount of data collection required, the cost is sometimes estimated in the number of gigabytes that are collected and processed.

Q: What can you recommend for certification for those who will be working in our organization's digital investigative and/or electronic discovery facility?

A: This topic is always a subject of intense debate. Currently there is no defined standard certification that is automatically recognized as being the authority on digital forensics. The federal law enforcement and military investigative agencies have developed formalized training and certification requirements for law enforcement and military personnel that conduct digital investigations, but unfortunately, there are no similar standards in the corporate sector.

Q: Our enterprise has a small digital investigative and electronic discovery capability. Should we be concerned with accreditation of this facility?

A: In the same manner that there has been little consensus on a certification standard for individuals that conduct digital investigations and/or electronic discovery, there has not been much agreement on accreditation standards. The most widely used accreditation standards from the American Society of Crime Laboratory Directors (ASCLD) are extremely rigorous. ASCLD has an accreditation standard for laboratories that process digital evidence. Enterprises that have met the stringent requirements, which in some cases are a multiyear effort, tend to be law enforcement agencies.

Q: Would it not be easier to build a team that conducts digital investigations before building the capacity to provide electronic discovery services?

A: If your enterprise will also be involved in legal proceedings, it certainly will take more time, effort, and money to build the digital investigative and electronic discovery capability at the same time, but because the policies and methodology utilized in this type of work are so similar, it makes the most business sense the develop it at the same time. The economies of scale that will be realized will more than offset the initial issues.

Chapter 4

Integrating a Quality Assurance Program in a Digital Forensic Laboratory

Solutions in this chapter:

- **Quality Planning, Quality Reviews, and Continuous Quality Improvement**
- **Other Challenges: Ownership, Responsibility and Authority**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

Over the course of my 30 plus years in law enforcement and forensic science, I have been involved in various aspects of quality systems development. I have been exposed to "Management by Objectives," "Total Quality Management," plus many of the latest business management fads. While each had elements of interest and value, none really struck a chord with me. It was not until I participated in formal project management training and worked in a "projectized" organization that I internalized the true value of an effective quality management system.

Quality systems are not an independent component of a work environment or project. They are part of the overall operation and product of a business unit, seamlessly integrated into the business processes and daily operations. I have come to embrace the concepts of W. Edwards Deming, Walter Shewhart, and Joseph Juran—that is, quality management processes, statistical sampling, and trend analysis form the basis for driving out product error and improving the efficiency and effectiveness of an organization's processes.

An effective quality management system is woven into the fabric of the operational and administrative business components. It impacts organizational planning, resource management, scheduling, and risk management to facilitate a proactive, upward and endless spiral of continuous quality improvement. Quality systems have no finish line. Quality is not a ticket to be punched. It is the organization's methodology to meet client expectations, needs, and to stretch the organization to adapt to changes in client needs, and to strive for a higher level of excellence.

This concept is applicable to any activity or business product, whether putting together a bicycle for your child on Christmas Eve or operating a Forensic Laboratory. It is built on the foundation of understanding your client's requirements and expectations, and then meeting or exceeding those expectations. When assembling the bicycle, you follow the process (directions) for putting the bicycle together to ensure that it is safely and correctly assembled to your child's needs and expectations to ride the bike immediately on Christmas morning. You may recheck the nuts and bolts for tightness, gauge the tire pressure, and maybe take it for a spin so that it meets the needs of your child the following morning. Your client, whether the judicial system or private sector, has similar expectations of quality in forensic processes, reports, and expert opinions—that your analysis can be relied upon as part of the judicial process or corporate goal as being derived from measurable, repeatable, and validated processes.

- How is a quality system implemented in a digital forensic laboratory?
- What are the benefits of an effective quality system?
- What resistance will I encounter when implementing a quality system in a laboratory?

This chapter will attempt to provide a framework based upon my experience of implementing an effective quality system in a digital forensic laboratory and answer the questions posed above, as well as address some of the challenges to implementing a quality system in existing organizations. My goal is that you will recognize the value of a seamlessly integrated quality system that supports the cycle of continuous quality improvement.

Quality Planning, Quality Reviews, and Continuous Quality Improvement

Before we can discuss methods for implementing a quality management system in a digital forensic environment, it is important to define a baseline understanding for the term "quality."

The American Society for Quality (ASQ: www.asq.org) defines quality as:

> A subjective term for which each person or sector has its own definition. In technical usage, quality can have two meanings: 1. The characteristics of a product or service that bear on its ability to satisfy stated or implied needs; 2. A product of service free of deficiencies....

The ASQ definition is exceptionally relevant to forensic science relative to client's needs and the standard for deficiencies—that is, "free of deficiencies."

Deficiencies and Driving Out Error

Let us first discuss the phrase "free of deficiencies."

In some business sectors, product deficiencies are an accepted part of business operations or manufacturing. For example, the food industry has acceptable levels of food contamination for human consumption. The U.S. Food and Drug Administration issued *The Food Defect Action Levels*, which defines the maximum levels of "natural or unavoidable defects in foods for human use that present no health hazard." This sets a standard for acceptable deficiencies in food products, for example, the level of acceptable insect parts in peanut butter. While I would prefer to have peanut butter that is completely free of deficiencies, it is not practical (in the FDA's opinion) to meet that level of quality, since peanuts are grown outdoors and a certain level of insect parts will inevitably enter the product during the manufacturing process.

Many businesses operate without any formal quality system. They may evaluate the cost of implementing a quality program against the costs of deficiencies in a decision matrix to produce a cost/benefit value of a quality system. Deficiencies are tolerated as part of the overall manufacturing, shipping, or other activities as part of a business' overall operations plan with acceptance of the fact that a percentage of products, a relative small, hopefully, will be produced that are defective. This may be referred to as a cost of noncompliance.

Recently, I ordered a lamppost and light fixture from an online supplier. About a week later, two boxes arrived. One contained the lamppost. The second contained wall art. In reviewing my order, I found that I had accurately ordered the post and fixture. However, I noticed that the fixture and the wall art had similar product codes—in fact, there was only one digit difference. I contacted customer service. I was informed that the fixture would be delivered in a few days. I asked about returning the wall art. I was told to keep it. Why? The cost of shipping and restocking it at the warehouse outweighed the value of the art. The mistake in reading the product number cost the company about $100, which seemed to be an acceptable loss or level of noncompliance as part of doing business.

Is there more than a loss of $100 in this example? Yes; the vendor now has incurred use of additional personnel resources to correct the mistake (customer service) and rework in filling and shipping the order (warehouse and shipping). I was impacted in terms of schedule. I had planned to install the lamppost during the upcoming weekend. Now, the installation would have to be delayed a week or more.

What about cost of reputation? Certainly, my opinion of this company was less than optimal. Although their prices were reasonable, just a few months earlier an order I received was missing a critical part causing completion of my installation to be delayed until the part arrived. The impact? Additional attention by customer service, warehouse pulling another unit, shipping cost, my schedule, and a diminished reputation for meeting my needs. This may be an acceptable loss to the company, however. And while I have a dim view of the quality of their warehouse processes, their prices continue to get me back as a customer because I am willing to sacrifice a little bit of time for money.

Another example of cost of deficiency was an order my wife and I placed for a name-brand entertainment center. The brand also was associated with a celebrity-named product line. Upon arrival, there was damage to the wood on the front. Not significant, but enough to file a claim for repair. Within a few days, customer service contacted me to advise a new unit would be shipped rather than repairing the existing unit. The new unit arrived within a few days. I asked the freight driver about taking the old unit and was advised that he had no paperwork to retrieve the unit. I contact customer service and left a message. To date, no one has ever called to retrieve this $800 item of furniture. The second unit also had minor damage which I decided to repair myself. Loss: Product cost, employee's performing re-work, shipping cost, and my future business.

While these concepts of noncompliance in the delivery of products may be acceptable in the manufacturing, merchandising and shipping industries, can a business engaged in providing digital forensic services afford such deficiencies in client services? Absolutely not!

Let's look at the Houston Police Department Laboratory (HPD-LAB) issue. In 2002, the HPD-LAB's DNA/Serology Section came under scrutiny resulting in discontinuance of those services. A preliminary report issued by the independent investigating team (www.hpalabinvestigation.org) cited lack of quality control and quality assurance as one of four major issues upon which the team was focusing. The team's fifth report, issued in May 2006,

identified 43 DNA cases and 50 serology cases where reliability and validity of the results and conclusions raised significant doubts. Although many factors led to HPD-LAB's critical failures, noncompliance with their Quality Management System, which was also cited as lacking relevance, was a major factor in their failed operations. While the jury is still out on the total cost, I suggest the following as having a negative impact on the organization in terms of the cost of noncompliance:

- Cost: Taxpayers are footing the bill for the investigation, rework, and contracting for outside DNA examination services.
- Schedule: Reexamination of cases may cause delays in the administration of justice.
- Reputation: HPD-LAB's reputation was certainly impacted and the impact may be perceived negatively across the organization and into the parent organization (HPD).

Nearly four years later, with millions spent, the HPD-LAB has instituted many changes, including implementing a relevant quality management system. The DNA lab recently received accreditation for one-year (as opposed to the usual five year) from the American Society of Crime Laboratory Directors-Laboratory Accreditation Board (ASCLD-LAB), which will allow them to re-institute services.

NOTE

Several states require laboratories to be accredited in order to perform forensic services. Texas and New York are two such states.

Is it possible for your product to be free of deficiencies? Although even highly evolved quality models, such as Six Sigma, set a target standard for acceptable levels of defects in manufacturing (3.4 per million), instituting an effective quality system can significantly reduce defects and drive out the possibility of a critical error in your digital forensic product for the client.

A quality system defines requirements for a digital forensic laboratory operation and its personnel to perform standardized processes that produce a consistent product. For example, examiners may be required to paginate, initial, and uniquely identify each page of their examination notes. This is done to ensure that the processes performed on a particular engagement are complete, organized, and properly associated with the engagement and processes performed on a computer (and for other business reasons you decide are important). During a quality system audit (discussed later in the chapter), it may be determined that one page of the notes were not initialed by the examiner. A review of case notes by the examiners

and other examiners indicate no trend of missing initials across the organization. It was a mistake. The remediation is to simply return the noncomplying documentation to the examiner for correction and remind other personnel to be mindful of the quality system's requirements.

What risk did this pose to our core mission and output to clients, which is examination processes, analysis, conclusions, and meeting our client's needs? I submit, none. The technical findings are the same and met the standards of being verifiable and reproducible. The output to the client still meets the client's needs and expectations for quality.

One key to developing an effective quality system in a digital forensic laboratory is to focus upon the client expectations, output, and examination processes to eliminate critical or fatal errors. How is this accomplished?

Let's continue.

Meeting Client Stated and Implied Needs

The challenge for the forensic laboratory in an emerging forensic field, such as digital forensic analysis, is that it is may be difficult to define quality in terms of client stated and implied needs. A client's needs vary from exam-to-exam or engagement-to-engagement; a client's abilities to digest the complexities of digital data storage differ greatly; and the intricacies of digital device hardware and software across many platforms creates a complex matrix of forensic analysis and digital data analysis output.

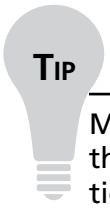
There are nuances of how data is stored from software manufacturer to manufacturer, and sometimes from version to version. The sheer number of digital devices and how the devices physically and logically store data (computer versus removable media versus cell phones/personal data assistants, for example) and the varied software used by different devices is a great challenge. Add to the mix the many methods people actually use devices (for example, iPod as a storage container for stolen trade secrets or stolen identities) and a rapidly changing technology environment, at best a moving target, and we are left with a daunting task of adequately and accurately defining client needs.

Our clients have different levels of expectations. Some are time sensitive like searching for a missing child or preventing a terrorist attack. Others may focus upon minimizing corporate financial exposure when the company's computer system is penetrated.

Quality is achieved through a process of understanding the client's needs and defining the scope of the engagement in terms of processes, schedule, and end product. As digital forensic service providers, we must also understand the implied needs of the client. Certainly, the client expects that we are using measurable, repeatable, and verifiable methodologies to perform our services, which are implied needs. For the law enforcement client, those implied needs may also include that the examination approach and outputs are focused by the specific laws alleged to have been violated and that the examination will not exceed the scope of the legal authority to conduct the examination (search warrant, court order). For example, a client investigating a financial fraud would have an implied need for a product

output that focuses upon financial transactions and included recovery and analysis of spreadsheets, financial transactions, or other computer use that are case-specific relating to the alleged violation. Similarly, in the private sector, generally accepted best accounting practices, records management, and regulatory compliance requirements provide a framework to meet the minimum implied needs of the client.

In order to achieve client needs and meet the expectation of quality output, examiners must also gauge their client's understanding of the complexities of digital forensic analysis and manage the client's expectations of the technical challenges posed in our discipline. What is it the client believes may exist in the digital device and how do they desire the analysis output? This is the starting point for establishing the required baseline analysis and, more importantly, setting expectations and defining scope of work to be accomplished for the client.



TIP

Many clients requesting digital forensic analysis services have little concept of the process for examining digital evidence, the volume of data the examination could produce, and the time required to conduct such an analysis. Before beginning your examination, always discuss the case particulars and expectations of the client. Manage expectations to reality. Do not leave the clients with the expectation that there is a "Solve Case" function key on our forensic examination machines as suggested by the CSI television programs.

Tools & Traps…

Beware of the client who requests "everything" or "just a quick look"

While client needs are always considered, managing expectations or guiding clients is always required.

If you give the "I want everything" client what they ask for, they most likely will be overwhelmed by the volume of data. Ultimately they will be back to you for help, dissatisfied with ***your*** product output. Manage those who try to entice you into a "quickie" examination. My experience is those are the examinations that end up taking the longest and run the highest degree of risk for examination error and standard operation procedures (SOP) noncompliance. When confronted with this type of client request, I think to myself, ***"There's always enough time to do it over, but never enough time to do it right!"***

While meeting client needs and quality standards is built upon interaction and documenting the scope of work to be performed, your organization must have a defined place where all work begins and the minimum processes that will be performed in every examination using measurable, repeatable, and verifiable procedures. These standards are detailed in your organizations Standard Operating Procedures and Quality Assurance Manuals, discussed later in the chapter.

Continuous Quality Improvement

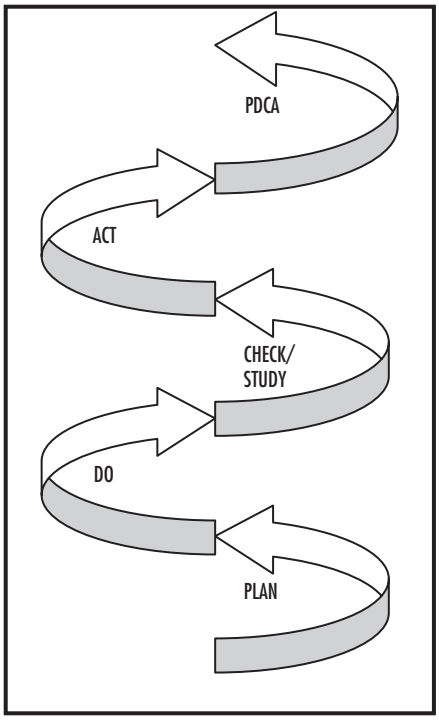
The International Organization for Standardization (ISO: www.iso.org) has developed the ISO 9001:2000 international standards for establishing quality management systems, which discusses the eight quality management principles:

1. Customer focus
2. Leadership
3. Involvement of people
4. Process approach
5. System approach to management
6. Continual improvement
7. Factual approach to decision making
8. Mutually beneficial supplier relationships

I have briefly touched on customer focus. Let's discuss ISO's comments on principle 6, "Continual improvement of the organization's overall performance should be an objective of the organization," through a concept of continuous quality improvement which will embody principles 1 through 6 above.

How does a digital forensic laboratory establish a methodology for continuous quality improvement? Allow me to share my implementation strategy for continuous quality improvement. I use the Deming Model adapted from Walter Shewhart, referred to as the PDCA or PDSA (Plan-Do-Check-Act or Plan-Do-Study-Act) cycle as shown in Figure 4.1.

Figure 4.1. Plan—Do—Check—Act[1].



The Plan-Do-Check-Act (PDCA) or Plan-Do-Study-Act (PDSA) model provides a framework or process for continuous quality improvement. It is a continuous process that spirals upward dynamically to the changing needs of the client in response to changes in client mission and/or forensic technology methodologies. This model is also exceptionally adaptable to building a quality system for general laboratory operations, which will be discussed under the Quality Assurance and Quality Control sections of this chapter.

Laboratory Planning

In new organizations, building a partnership across the organization can be accomplished through the development of a concept of operations (CONOPS) or business plan (I will use businessplan or BP to describe both) that integrates the business strategy into a design plan. In new organizations where new services are planned, buy-in can be garnered from top to bottom by seeking input from across the planned organization or setting expectations during the start-up or initiation phase by having new employees review and comment on the BP.

Because digital forensic laboratories are a relatively recent addition to the forensic science community, using a BP to describe the organization prior to it becoming operational may be critical for employee orientation and managing expectations of the parent organization or executive management, particularly when the plan includes a quality system for a digital forensic laboratory. As such, the BP becomes instrumental in garnering both executive support for adequate funding and personnel resources toward a successful and attainable implementation as well as employee support of the BP roadmap. The BP should define the overall laboratory objectives, mission, scope, organization, and general characteristics of the

laboratory administration, quality systems, and operations. Most importantly, the BP forms a defined foundation for a clear vision of operations that are relevant and achievable.

Existing laboratories may draft a BP for implementation of digital forensic services when the services are a new component to the laboratory. Certainly, many of the overall laboratory characteristics would be readily incorporated into the BP. However, the BP would focus on those administrative and operation aspects of the new service, aspects of the new division that may require specific adjustments to existing operations, and how it would be integrated in the existing organizational framework.

Probably most challenging is instituting a plan that represents a significant change to the methodologies already in place for providing digital forensic services. Why? In the law enforcement sector, for example, until just a few years ago nearly all digital forensics were conducted as an investigative methodology by sworn officers who also conducted the investigation. The vast majority of these investigative units have neither documented standard operating procedures nor quality management systems for their digital forensic services.

In the private sector, existing services are part of an investigative or audit team. The digital forensic personnel are considered a resource to the team to accomplish the contracted task. The team may not view digital forensic tools and methodologies as nothing more than a tool of the trade requiring no application of forensic science principles. When I raised the scientific principles of using measurable, repeatable, and verifiable processes to a government attorney recently, he responded, "…it [digital forensics] is not DNA." No, it obviously is not, but should the same principles apply?

In 2003, digital forensics was recognized by the scientific community as a forensic science discipline. ASCLD-LAB defined specific accreditation criteria for a digital forensic laboratory in ASCLD-LAB's 2005 standards titled, Digital and Multimedia Evidence with sub-disciplines in Computers Forensics, Audio, Video, and Image Analysis. At this writing, 14 laboratories have achieved ASCLD-LAB accreditation in Digital and Multimedia Evidence. Our laboratory, the New Jersey Regional Computer Forensic Laboratory (NJRCFL), was the fifth in the United States to be accredited in both computer and video sub-disciplines just short of our second anniversary. There were many keys to our rapid success, but it all began with a jointly-written (management and examiners) CONOP that clearly defined the mission, structure, and services that were relevant and achievable.

I submit that a BP for a digital forensic laboratory must acknowledge that digital forensics is recognized in the scientific community as a forensic science discipline. As such, the plan must include implementation of a quality system as well as a plan for accreditation.

The next step is to define how the mission, goals and services are to be specifically implemented through standardized baseline processes. standard operating procedures (SOPs) and quality assurance manuals (QAMs) are essential for a digital forensic laboratory. They define the minimum baseline procedures for conducting a forensic analysis as well as the quality system components to assure client's needs are met and drive out the risk of critical error in the output.

The Structure of an Organization's SOPs or QAMs

Here is an example of the structure and topics that may be included in your organization's SOPs or QAMs.

Mission Objectives

- What does the laboratory want to accomplish
- Core laboratory purpose, client satisfaction, and business philosophy
- Type and Extent of Services

Laboratory Administration

- Location and Contact Information
- Personnel Guidelines and Regulations

  Work Hours

  Leave Policies

  Dress Code

  Grievance Policy
- Media Policy
- Health and Safety Program

Personnel Responsibilities and Job Descriptions

- Organizational Chart
- Training Program
- Examiner Qualifications and Duties

  Performance Evaluations

  Continuing Education Requirements

  Competency Testing

  Proficiency Testing
- Management Responsibilities
- Ethics and Confidentiality

Laboratory Operations

- Services: Field or On-site Services

  Laboratory Services

  Consulting Services
- Examination Documentation
- Workflow
- Records Management
- Evidence Management

  Chain of Custody

  Evidence Storage

Standard Operating Procedures

Standard operating procedures include hardware- and software-specific technical processes. These processes address the following areas:

- Scope/Purpose of use
- Limitations of use
- Procedure for use
- Calibration
- References
- Notes
- History
- Approval

Testing and Validation Procedures are also elements of standard operating procedures.

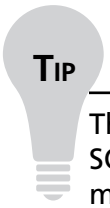
Quality System

- Management and Staff Commitment to Quality
- Controlled Documents
- Quality Planning and Policy
- Technical and Administrative Reviews
- Annual Audits and Annual Quality System Review

- Corrective Actions
- Monitoring Examiner Testimony
- External Complaints and Resolution

Document History and Approval

Continuing as defined by your organizations goals, mission, and client services…



TIP

The components outlined in the preceding section are not all inclusive. Your SOP and QAM components will be driven by your organization's specific mission, goals, services, and methodologies to meet client expectations.

The SOP and QAM are often two separate documents.

Does your organization have a goal to become an accredited facility? If so, a good starting point for your QAM would be to be consistent with the format of the accrediting body's criteria. This sets the baseline for compliance against required accreditation criteria.

Essential components of the planning process are embodied in principles 2 and 3 of ISO 9001:2000.

Principle 2 (Leadership) reads:

> "Leaders establish unity of purpose and direction of the organization. They should create and maintain the internal environment in which people can become fully involved in achieving the organization's objectives."

Principle 3 (Involvement of people) reads:

> "People at all levels are the essence of an organization and their full involvement enables their abilities to be used for the organization's benefit."

As a laboratory director, my role has been to provide the benefit of my experience in law enforcement, management, and education to guide the development of the organization through my staff. I describe my role as one of facilitator and leader. I use the following analogy to describe my implementation of principles 2 and 3:

A poor manager watches the subordinates decorate the office Christmas tree. When it is time to place the star on the tree, the poor manager places the star atop the tree, stands back, and announces to all, "Look at the beautiful Christmas tree I decorated."

An effective leader asks his work family for their opinion on the ornaments to acquire for the decorating, advice on which ornaments shall be hung in the appropriate sequence, participates and listens to suggestions for ornament adjustment, schedules sufficient time to accomplish the decorating process to be ready for the office party, and facilitates the placing of the star by holding the ladder. At the office party, the effective leader recognizes everyone for their joint accomplishment, and all take pride and enjoyment in the finished product.

Organizational involvement in the planning process that embodies buy-in from top to bottom will result in an effective plan that can be seamlessly integrated into the organization's business processes. As depicted in the PDCA illustration, the plan cycle will be revisited each full PDCA cycle. Thus, it is critical, in my opinion, that the plan be organically grown within the organization so that ownership in the plan and its review cycle are embodied as a component of the laboratory and its personnel.

"Do" or Executing the Plan

Once the plan is established, elements of the plan are executed by applying the resources designated—people, equipment, software, SOPs, quality systems, and support components.

In the "Do" phase, the laboratory is performing services defined by the mission and goals. We are providing client services. In addition to providing services during this phase, we are also gathering data about the client's assessment of our services and documenting our own internal assessments of operations and services. The gathering of data is part of the continuous quality improvement process and includes activities of quality systems termed "quality assurance" and "quality control." What are quality assurance and quality control?

The Project Management Institute (PMI: www.pmi.org) defines *quality assurance* as:

> Applying the planned, systematic quality activities to ensure that the project employs all processes needed to meet requirements (client needs).

ASQ defines quality qssurance as:

> All the planned and systematic activities implemented within the quality system that can be demonstrated to provide confidence that a product or service will fulfill requirements for quality (client's needs).

ISO defines quality assurance as:

> Those planned and systematic actions necessary to provide sufficient confidence that a laboratory's product or service will satisfy given requirements for quality.

The common elements across each definition that are essential to digital forensic laboratory operations are "planned and systematic activities to produce output that meets the client's needs." The client's needs were discussed above. What are the planned and systematic activities? They include such elements as adherence to standard operating

procedures or tested and validated methodologies that produce repeatable, measurable and verifiable results. An important component of a quality assurance qrogram is conducting Quality audits to identify ineffective policies and processes. Quality audits are one major example of planned and systematic activities of a quality management system. The outcomes of quality audits are:

1. *Confirm* processes that produce the intended results
2. *Identify* weak or ineffective processes that impact laboratory administration or operations
3. *Initiate* corrective action and remediation
4. *Document* lessons learned

PMI defines *quality control* as:

> Monitoring specific project results to determine whether they comply with relevant quality standards and identifying ways to eliminate causes of unsatisfactory performance.

ASQ defines quality control as:

> The operational techniques and activities used to fulfill requirements for quality.

ASQ acknowledges that the terms "quality assurance" and "quality control" are often used interchangeably.

ISO defines quality control as:

> Internal activities, or activities conducted according to externally established standards, used to monitor the quality of analytical data and to ensure that it satisfies specified criteria.

Clearly the common elements across each definition essential to digital forensic laboratory operations are techniques and activities to ensure that the client's needs are met, and to eliminate causes of error or unsatisfactory performance.

While we are performing client services during the "Do" phase, we are also gathering data about our quality system and the quality of our services by implementing within the quality system planned and systematic activities to ensure we are meeting the client's needs and eliminate causes of error or unsatisfactory performance. Without effective quality components, we elevate the risk of noncompliance and significant error that the FBI Laboratory experienced in the 1990s.

The Department of Justice Office of Inspector General's (OIG) April 1997 report titled, *The FBI Laboratory: An Investigation into Laboratory Practices and Alleged Misconduct in Explosives-Related and Other Cases* (April 1997), listed 12 recommendations to enhance the quality of

the FBI Laboratory besieged by allegations of examiner misconduct and poor operations/administration. In my opinion, 11 of those 12 recommendations are directly related to basic quality assurance and quality control criteria. They include:

1. Attain accreditation from ASCLD-LAB
2. Defining the roles of laboratory examiners and resolutions of disputes
3. Report preparation based upon evidence analyzed by each examiner
4. Peer review (technical review)
5. Case documentation that is complete with notes, worksheets, and other documents upon which examiners used to support conclusions
6. Record retention
7. Examiner training and qualifications with uniform curriculum and moot court
8. Examiner testimony monitoring
9. Examination protocols documented
10. Evidence handling protocols
11. Defining the role of management in articulating the laboratory's vision, goals, and priorities

NOTE

A 12th recommendation has to do with restructuring of a forensic unit.

Each of these recommendations, effectively implemented, will improve quality and drive out error. In the FBI's case, the OIG report, dated June 1998, articulated the implementation of the above list quality criteria. The FBI Laboratory attained accreditation in 1998 and has retained accreditation since.

How is quality control implemented? One example of a method for implementing quality control is to establish configuration control for documentation. Creating a baseline template for laboratory reports with required data entry fields for examiner entry may prevent the report from being finalized and sent to the client without critical data. Another method is establishing configuration control for standardized hardware and software to be used in each baseline examination. That baseline would include only the tested, validated, and approved software applications for conducting an examination. A third example of quality control is implementing a process of required technical (peer) and administrative

reviews of examination output prior to clients receiving the product to ensure that the examination procedures document are repeatable, verifiable, and results are validated, as well as confirming that applicable components of the quality system have been met (for example, are the exam notes initialed).

How is quality assurance implemented? Quality audits or reviewing components/processes of laboratory operations critical to client services for compliance (or noncompliance). I mentioned the example of an audit of case documentation above where closed cases are reviewed for noncompliance from both a technical approach to administrative quality standards. Another component of the quality audits includes a review of the evidence storage processes and documentation. This includes review of "chain of custody" documentation as well as physical access documentation to the evidence vault, for example.

The specifics of quality control and assurance for your organization will be driven by your mission, goals, and perhaps statutes, regulatory requirements, and client expectations.

In summary, the execution or "Do" phase is where the weaving of the laboratory's forensic processes with the quality system occurs with the goal of both meeting or exceeding the client's needs while gathering data that will facilitate an evaluation of the quality system's effectiveness.

"Check" or Study Processes

Properly implemented, a quality management system will provide data about your internal quality processes and whether the forensic services are meeting the client's needs and expectations. While data about the processes that include quality assurance and quality control is being gathered during the "Do" phase, the analysis of the data and planned remediation or adjustments to the process are a function of the "Check" phase.

I discussed several examples of internal audits of the quality system that include a physical review of a percentage of examination case files or client files, and related chain of custody documents, client requests, reports, examination notes, and documentation to gauge compliance with internal and external operating procedures and quality policies. The audit looks for trends of noncompliance, which may suggest flaws in the organizations processes.

For example, an internal audit of the examination case file documentation observed a pattern of inconsistent administrative review "check-offs" against the question, "Is the Evidence Chain of Custody (CoC) accurate?". Some reviewers marked "NA" (not applicable) while other reviewers entered a checkmark in the box. Digging further, we find that the reviewers entering "NA" have interpreted the question "as having reviewed a full CoC detailing every movement of the evidence from the time the evidence came into the laboratory until its current state which was not filed with the case documentation." The reviewers entering a checkmark in the box interpreted that the existence of a receipt for the evidence with accurate data (what, when, where) was sufficient to meet the review criteria. The audit of the quality system for case documentation cites the inconsistent responses, identifying two

flaws. One flaw was in the wording of the required "check-off" and the second that the reviewer's had not been provided training in the meaning of the question, both being process flaws. Remediation: Adjust the wording and educate the reviewers of the wording's intent.

I am a firm believer in having a lessons-learned process. The process should involve some formalized output, documentation of the issue, resolution, and sharing. This is sometimes formalized through the audit and corrective action process. However, there are challenges and issues that arise that may not rise to the level of a formal corrective action, SOP, or QAM provision. Occasionally, it is simply a management discretionary matter that improves process or efficiency. It also may be a slight change in examination methodology that creates a better output that does not rise to the level of an SOP requirement, but simply enhances the output to the client. The lessons-learned library could be an excellent reference when revisiting of the "Check" phase.

It is essential that the laboratory create a centralized library of lessons learned from these events. Here are a few ideas for implementing the lessons learned component:

- Blog or information-sharing bulletin board on your agency's Intranet
- E-mail to your personnel (be careful of information overload)
- Searchable electronic library service
- Laboratory meetings
- Paper documentation file

Audits and activities during the "Check" phase should not be punitive. This must be conveyed to organizational personnel. It is essential that the "Check" phase include input from all levels of the organization involved in the processes being reviewed, including actual participation of the examiners in audits, lessons learned, recommendations for remediation or tweaks in the processes, and modifications to SOPs and QAMs.

"Act" or Adapt and Refine the Plan

In the example from the prior section, a corrective action request might be written against the inconsistent administrative review relative and a remediation plan documented. The plan might simply result in a wording change to the check-list and the education of the reviewers to bring a common understanding and application of the review criteria.

The very nature of our business is constantly changing and requires adaptability. However, I caution you not to shoot from the hip. Measuring the effectiveness and efficiency of a laboratory's quality systems requires documentation against which to measure. Defining that system against which to measure requires a careful aim at the target which is meeting or exceeding client expectations in what seems to be a rapidly changing environment. In the field of digital forensics, clients' needs change based on the type of industry, purposes of the engagement, compliance with new regulatory requirements or laws, and other case-specific

factors. This is compounded in our industry by the rapid changes in technology that challenge our technical approach and test our resolve to complete the mission. The bricks seem to always be moving under out feet!

An effective implementation of a quality system means that it is extensible—, that is, your organization must understand organically that changes in processes are inevitable. The "Act" phase is where your laboratory adapts processes to meet client expectations in our rapidly changing forensic science environment.

Continuous Upward Spiral of Excellence

How long is each phase of the PDCA cycle? My experience is that the full cycle be no less than once a year and be part of the annual audit and review of the quality system. However, in our changing technological environment, it is sometimes necessary to review at least some elements of the plan more often. However, I would recommend no more than every six months for major change.

Tools & Traps…

Too Much Change!

When starting a new laboratory for digital forensic services, many processes may be new in terms of facilities, administration, and operations. In law enforcement labs, some examiners may be police officers who are conducting forensic exams as part of their investigation with no SOPs or QAMs. There is a level of frustration with the move from investigator role to forensic examiner.

One methodology is to reduce frustration and to obtain buy-in of the quality system by avoiding the "big bang" approach to change. "Spoon feed" the quality system in an existing organization to minimize the impact on operations (examination output) and, more importantly, morale.

Cost of Quality: Why Bother?

I have heard it said that accreditation is the latest fad. "Suck it up until we are inspected and accredited. After that, things will return to normal." Nothing could be farther from reality.

ASCLD-LAB's accreditation process, whether Legacy or International, states the following objectives for laboratory accreditation:

- To *improve* the quality of laboratory *services*.
- To adopt, develop, and maintain standards that may be used by a laboratory to *assess its level of performance* and to *strengthen its operation*.
- To *provide an independent, impartial, and objective system* by which laboratories may benefit from a total operational review.
- To offer to the general public and to users of laboratory services a means of identifying those laboratories that have demonstrated compliance with established standards.

Accreditation of a digital laboratory is an independent assessment of your laboratory's quality system. Accreditation means you meet your standards for quality services to your client.

I had a conversation with a private laboratory director recently about his view on independent accreditation. Simply stated, his organization implemented a quality system that was compliant with accreditation standards but had no intention of seeking accreditation through an independent inspection. Why not seek independent validation of the system through accreditation? It reminded me of one of my friend's favorite sayings: "It's like wetting yourself in a dark suit. It gives you a warm feeling, but no one notices."

I mentioned this to my colleague Dr. Daniel Garner, Chief of Forensic Science, International Criminal Investigative Training Assistance Program, Department of Justice, which generated a discussion of the value of implementing a quality system in a forensic laboratory as well as attaining industry-wide recognition of an independent inspection of the system. Dan was formerly President of Cellmark Diagnostics, Inc., a DNA laboratory. Under Dan's leadership, Cellmark was the first private laboratory in the United States accredited by ASCLD-LAB in DNA. Cellmark was noted for having the highest quality services as well as a high-end pricing schedule "for which we apologized to no one," according to Dan. Prior to being President of Cellmark, Dan also ran the Alcohol, Tobacco, and Firearms (ATF) national laboratory in the Washington, DC area, where he was successful in attaining the first federal laboratory accreditation.

While running Cellmark, Dan had a contract to provide DNA analysis services to the City of Los Angeles. On June 12, 1994, Nicole Brown Simpson and Ronald Goldman were murdered in the Brentwood-area of Los Angeles. Police found a bloody glove at the crime scene, much publicized during the trial as a critical item of evidence. Dan recalled competitors offering to conduct the DNA analysis of the glove for free. However, the City of Los Angeles opted to stay the course with Cellmark because of their high quality service and reputation in the forensic science community.

Dan and I traveled to the Ukraine in 2006 at the request of the Ukraine Ministry of Interior and the U.S. Embassy to assess the Ukraine's national forensic laboratory. The ministry had set a goal of achieving international accreditation under ISO 17025 for the

laboratory, which included both a DNA unit and a digital forensic unit. While using somewhat dated techniques in DNA and digital forensics, everyone, from the "experts" conducting examinations to the top-level executive management, recognized the importance of seeking independent accreditation. They understood that accreditation meant that their quality system and standard operating procedures met the standards and best practices of their forensic science disciplines against international criteria. They desired recognition among their peers. They wanted to stand should-to-shoulder with their accredited peers as having met the standards for excellence.

In the private sector, independent accreditation could mean standing above your peers as Cellmark had. If your clients had a choice between a laboratory that is accredited and one that is not, which do you believe they would choose? In my lamppost example above, I continued to stay with a supplier that seemed to make frequent mistakes in order to save money. When it comes to digital forensics, your data, your business's reputation, and serving the judicial system, should your organization strive for an exceptional level of excellence? I know I would and did!

The quality cycle never ends. Quality improvement is never finished and is not a medal that is worn on the chest. While my laboratory proudly displays our ASCLD-LAB accreditation certification and the ASCLD-LAB seal, my personnel were doing the same processes before, during, and after the inspection while making adjustments, adaptations, and refinements to our quality system. While the illustration above only depicts one cycle of the PDCA/PDSA process, this cycle continues to push the organization to a higher level of achievement.

Other Challenges: Ownership, Responsibility and Authority

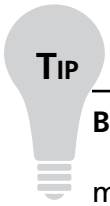
Ownership! That is the key to an effective quality system. Everyone must take ownership. Responsibility and accountability are essential components of ownership. Remember the Christmas tree analogy? The work family takes ownership by selecting the ornaments and working together to meet the goal. In the forensic laboratory, it is the same. Management must not simply use slogans and sound bites to show their support for quality processes; they must demonstrate their total commitment if they expect the employees to follow.

While contemplating opportunities post retirement, I had a series of interviews with a large company. Each interview except the last one centered on building (increasing) digital forensic services from a quality system foundation. At the last interview I was asked how long it would take before the company would realize a benefit from a quality system;—that is, increase its billable hours. Because I am a PMP, I always look to manage expectations. I responded that it would take two years, although I believed the benefit could be attained sooner (I did it in less). The quality system can be developed incrementally, which would likely bring incremental benefits to the organization and for clients. Ultimately, we did not

come to terms. One major factor was that I did not believe management had the fortitude to withstand the required investment to implement a quality system. Management must set the tone for ownership.

Ownership is established in several ways. I mentioned buy-in earlier in the chapter when discussing the CONOPS/Business Plan development. Another example in building ownership is to have the examiner personnel have a voice in the PDCA cycle. One method I mentioned earlier was configuration control. Establishing a configuration control board (CCB) staffed with examiners and an examiner as the chairperson provides an opportunity for organic development of new processes or refinement of existing processes. This establishes personal investment and ownership in the process for improving quality. From report and examination note templates, to baseline examination software configuration, the CCB takes ownership of change recommendations and analysis to improve examination processes and client output.

Together, the CCB examiners decide on the recommended approach for change, managing change implementation to no less than six-month cycles. Change to the processes should be reviewed by the laboratory quality manager for potential impact/conflict with the quality assurance system and with the laboratory director for impact on overall laboratory administration, regulatory requirements, and legal issues. Management review should be concluded efficiently to reflect the CCB's importance to the process and to ensure recommendations do not languish on a manager's desk, which could lead to frustration by the board and laboratory staff.



Before, During, and After the ASCLD-LAB Inspection—Ownership

In 2006, we were preparing for our ASCLD-LAB inspection. During a meeting with my staff, I referred to the inspection as "your inspection." I asked each person in the room to take a minute or two to advise their colleagues and peers what they have contributed to the quality system. Each staff member spoke about their commitment to quality through specific activities and methods. They accepted responsibility and accountability in their responses. They had vested ownership.

Management's Responsibility for Ownership in the Quality System

In my laboratory, I drafted the section of the QAM that sets forth management's (my) roles and responsibilities for taking ownership of the quality system. I asked for examiner feedback to help establish their expectations of me in taking the leadership role in the

quality system. Together, we defined our expectations for acceptance of responsibility by me and my executive staff (who also provided their views of management's quality system responsibilities and ownership). Together, we defined our expectations of each other in terms of responsibilities, authority, and performance.

The Quality Manager

Establishing the position of quality manager, who reports directly to the top management executive or laboratory director, is essential. The quality manager must be given both the authority and responsibility for managing the quality system and for speaking for executive management on quality programs. I have taken a few quality system "hits" from my quality manager when I have inadvertently stepped outside of the quality processes. But, the quality system is not intended to be punitive. It is part of the continuous quality improvement process. Management must communicate that its intent is to be a key component in the pursuit of excellence for employees, management, and the organization.

Tools & Traps…

Remediation Plans

While the quality system is meant to foster continuous quality improvement, it also sets expectations for remediation of deficient examiners. In my opinion, the system, not the employee, is usually to blame for most process errors with remediation of lab-wide processes; occasionally the worker is solely responsible for deficiencies. Remediation is usually the required response to deficient personnel. Counseling, additional training, closer review of work product, assigning less challenging examinations, and setting specific milestones for improvement are all potential components of an effective remediation plan. However, when these measures do not work, the employee may be dismissed because a forensic laboratory cannot afford to risk a fatal error in an examination that would impact its accreditation and reputation!

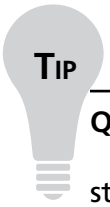
This is a slippery slope because it is the nonpunitive quality system that triggers the remediation event for the employee. Management must make it clear that quality systems are designed to improve processes and the overall product of the laboratory. However, employees are responsible for maintaining standards and performing at an acceptable level since we often hold the fate of an accused person or major civil litigation result in our work product.

Choosing a quality manager (QM) is important to success of the quality system, particularly focusing on the quality assurance and quality control aspects of the system. My first appointment of a QM was disappointing. I took a highly-respected examiner with a great deal of experience, thinking he was our best candidate in the laboratory to implement a quality system. Being a good examiner does not make a good quality manager.

After discussions with my managers, we selected Dave, a New Jersey State Trooper, for the position of QM. He is a bright, no-nonsense guy known for getting to the point. I have to admit, initially it concerned me. Coming from a paramilitary-style organization like the state police, Dave might not be a QM who could relate well to techno-geek free spirits. While I had come to know and view Dave's educational accomplishments and experience as exceptional, troopers by their very nature, training, and experience can be a bit too "forthcoming." I was building a system with a foundation of trust. I did not want it rammed down the staff's throats.

In a matter of several months, we incrementally instituted our quality system. While the laboratory worked together for ways to seamlessly weave the entire quality components into our workflow, Dave focused on implementing the auditing procedures to measure the effectiveness of the quality system. The quality system was producing meaningful documentation of our services and adherence to our quality standards facilitating an effective remediation of deficiencies or tweaking of the processes.

Now, nothing is forever and someday Dave may leave the lab for a promotion in another department of the New Jersey State Police. Eric, a pony-tailed geek analyst/examiner, came into my office one day and expressed his interest in assisting Dave with internal audits so he could learn more about quality systems. Eric by his nature is analytical and is exceptionally dedicated with a great deal of perseverance. All are very good traits for an effective QM. Continuity of Operations – Personnel! Save that for another chapter; however a good personnel continuity plan and an effective quality system will keep a lab functioning at a high-level of professional output during periods of critical personnel transition.



TIP

QM Responsibility/Authority

"My quality manager speaks for me on all quality issues," I state to my staff and anyone who asks about quality. If the QM has the responsibility, the QM must have the authority from the highest level of management.

NOTE

"Geek" is a very positive term and trait in our laboratory. And if you drink Mountain Dew, eat Twinkies, and work from noon until 2 A.M. or later, you are a "Super Geek."

Personalities and Patience

Perhaps this should be a book in and of itself?

My laboratory is unique—at least when comparing the attributes of laboratories outside of the FBI–Sponsored Regional Computer Forensic Laboratory (RCFL) program. We have examiners who are sworn officers and civilians from nine agencies. I estimate that we have somewhere around twelve different sets of work rules, three unions, coupled with each person's individual needs and personalities. Add to that the challenges mention above regarding moving digital forensics from an investigative responsibility to a scientific environment and patience can be stretched. What a mix to manage!

NOTE

As of this writing, more than 100 federal, state, and local agencies are participating agencies in the **Regional Computer Forensic Laboratory Program** in 14 laboratories nationwide. Go to www.rcfl.gov to learn more about the RCFL Program and how to become a participating agency.

This section most reminds me of Dr. Spencer Johnson's book, *Who Moved My Cheese*. In the late 1990s, some in the digital forensic community began to recognize a lack of consistency in the practices of digital forensics. It was not that evidence was being mishandled or improperly analyzed; it was simply the lack of consistency in the processes, including the data analysis, output, and reporting.

This recognition resulted in efforts from several organizations to formulate a baseline definition for the standardization of processes and outputs for each examination. Efforts in which I was involved were not intended to limit the art of digital forensics, but to establish guidelines that every examination should produce. From that starting point, the client's needs and the examiner's education and experience would drive the path of the examination and output. The Scientific Work Group on Digital Evidence (SWGDE: www.swgde.org) posted suggestions for baseline SOP and QAM documents, which include:

- Best Practices for Computer Forensics
- Recommended Guidelines for Developing Standard Operating Procedures
- Recommended Guidelines for Developing a Quality Management System
- SWGDE/Scientific Workgroup on Digital Imaging Technology (SWGDIT) Proficiency Test Guidelines

- SWGDE/SWGDIT Guidelines and Recommendations for Training in Digital and Multimedia Evidence
- SWGDE/SWGIT Digital and Multimedia Evidence Glossary

I have found that most successful examiners in digital forensics have common traits of perseverance, ingenuity, tenacity, and sometimes stubbornness. Most have found their own way in digital forensics dating years before the discipline was recognized as forensic science. Thus, they may perceive threats to their world as we move the discipline from investigations to forensic science and from a wide-open technical frontier to an environment with documentation requirements, quality systems, and baseline standardization.

My experience in law enforcement has been that the geeks conducting computer forensics were considered wizards performing magical feats of recovering deleted files, file fragments, and passwords. Management and investigators rarely understood how this wizardry was performed, nor did they truly care. But it was a great boost to their ego to be the go-to geek for computer cases!

For example, I recall a very seasoned FBI special agent who barely knew how to turn on a computer coming to me with a laptop that was suspected of sending an e-mail via America Online (AOL). I had to explain to him that Web-based mail was not like client-based email like Outlook Express. "The e-mail may not be on this computer, but let me do a few things and I will call you back," I explained. Within a few hours, I had recovered several versions of the e-mail from the hard drive's free space or unallocated space. The FBI Agent agent was surprised and exceptionally pleased when I provided him with printouts of the eMail e-mail message, which he used to confront the suspect, resulting in the subject confessing, and later pleading guilty.

In some cases, the examiners thought that standardization would limit the digital forensic examiner's ability to successfully perform and produce the results as they had in the past. I mentioned several times in this chapter that the standard operating procedures are the minimum baseline for conducting a forensic examination; they are not intended to limit the art the examiner brings through their technical and investigative experiences and education. It is essential that this be communicated.

Winning them over in my laboratory was one of several major challenges in building an effective, multi-agency discipline in digital evidence. New examiners were easy. They really didn't know the difference between "old" style forensics and what they were being taught relative to tools and methodologies detailed in the SOPs and the requirements of the QAM. Well, they didn't know better from experience, according to the "old fart" examiners. The challenge was convincing the old farts, who had been conducting forensics for several years, to give quality processes a chance. Their perception was that they would lose their independence and ability to apply ingenuity to problem solving and become "image monkeys" with no wizardly output.

I approached the challenge with two consistent lines of thought relating to the desire to continue performing "wizardry", by continuing to challenge my examiners to a higher form of technical challenge: (1) assess your client's needs, and (2) adapt to your client's needs. Let's discuss these two lines of thought.

Assess Your Client's Needs

Today, many clients are becoming computer savvy; but the range of technical challenges we are confronted with on a daily basis in digital forensics is exceptionally broad. No one can be an expert in everything. Certainly not the client nor those of us in digital forensics. Basic training and continuing education programs in law enforcement have included the importance of digital evidence and where that digital evidence is likely to be found. Training programs are abundantly available to law enforcement, many free of cost (for example FBI, Regional Computer Forensic Laboratory Program, Federal Law Enforcement Training Center, the National White Collar Crime Center (NW3C), and many others). In the private sector, companies offer free Web-based training seminars for attorneys and practitioners. As the level of expertise of our clients increases in the understanding of digital evidence, so will our demand for capacity in terms of volume of evidence items to be analyzed.

As computers and other digital devices become more commonly used across all aspects of life, and as law enforcement and private sector clients increase their familiarity with digital data challenges, digital forensic service demands are likely to be more technically refined or leading edge (sometimes bleeding edge). The expectations of our clients will shift from total dependence upon the examiner for all data recovery and analysis, to data recovery and assisting with understanding the most difficult technical issues of digital data storage and device activity, defeating encryption, and analyzing activity to provide circumstantial evidence (that is, timelines) on how to place the suspect at the scene of the crime. Remember, there is no "Solve Case" key on the keyboard and digital forensics always provides just one piece of the puzzle. The puzzle-piece definition comes from discussion with the client of the digital evidence as a tool or instrumentality of the crime, need for data to meet regulatory compliance, or other needs of the client that are required to be validated from an analysis of the digital evidence.

The examiner's challenge is assessing their client's level of understanding of digital evidence and providing a level of service that meets the need. For example, we have a number of very technically savvy investigator-clients who simply want the laboratory to preserve the evidence and provide an image for them to review and conduct the analysis. In fact, to accommodate this methodology, we have implemented a forensic storage area network that facilitates this methodology. It has greatly improved our output and permitted us to focus on more technically difficult forensic challenges and focus our client services on those needing more hand-holding. However, I caution the clients in their use of this methodology that, while an efficient methodology to interact with the data, look for information of lead value,

and to see how the digital evidence puzzle piece fits with the remaining evidence, that it is the examiner who should explain the technical aspects of digital data recovery methodologies and interpretation of the data and how it is eventually stored on a computer.

Adapt to Your Client's Needs

While your clients are increasing their understanding of digital evidence, there are often going to be very detailed technical questions and challenges that require in-depth knowledge of systems, hardware, and cross-platform issues requiring detailed technical analysis.

There are many stages to an investigation, audit, or incident response. As time moves along, the client's knowledge of facts and events continue to build a picture of the event under investigation. Later in the process, defense counsel or opposing corporate counsel may produce detailed, challenging digital evidence analysis reports that suggest alternate explanations for events than those proposed by your client based upon their review of evidence, including your analysis.

The challenge today is to be prepared to work with the client to refine their technical analysis requirements and to address the more complex technical issues. I liken the process to the intelligence cycle that is quite similar to the PDCA cycle. As information is collected, analysis of the information along with understanding of the information identifies gaps that result in "taskings" for additional investigation and research. Some taskings may come to the digital forensic examiner as part of refining and adapting to the changing information environment.

The level of technical analysis that is demanded from you will increase as users become more technically savvy and as we learn more from the intelligence cycle of the matter at hand. One example of the area for refinement by the examiner is when we are confronted with password-protected or encrypted documents/containers – a particularly challenging area. While some encryption algorithms may never be broken in our lifetime with current technology, other encryption can be addressed with success using tools readily available, but requiring a sound understanding of methodologies to efficiently employ. It may require a combination of your technical knowledge and the investigator's social engineering or tenacity for detail working together to "find" those passwords in the course of the investigation (this is more a methodology topic) that will help meet client's needs and expectations for the final product.

Private Sector Challenge

In the private sector, the examiner-level resistance to change may be present, but is less prevalent than in the public/law enforcement sector. The real challenge, in my view, is selling a formalized quality assurance plan and, perhaps, an accreditation plan to executive management where billable hours are not a component of the quality process. On its face, a quality assurance plan, as described above, may appear to be pure overhead to management.

I mentioned earlier my experience when considering private sector employment. That company was considering implementing a quality system in their digital forensic laboratories. Additionally, they were considering adding laboratory capacity in several foreign locations. This certainly added complexity to their stated desire to bring consistency and quality across the planned global operations. At the conclusion of the interview and negotiation process, the corporation believed the investment involved too much overhead without a return on the investment during the development period. Two years was too long. I believe that the private sector is avoiding the inevitable or at least an opportunity to stand above the rest.

How do you overcome this view of cost-impact when attempting to convince management of the value of an effective quality system? I point to the Cellmark example and benefits they derived as a positive return on investment (ROI) for a quality system. I also point to the HPD-LAB and FBI Lab's long road to rebuilding trust and reputation as a cost of noncompliance. Establishing a quality assurance system is absolutely essential to driving out risk of error regardless of whether your client base is the criminal justice system or private sector. Accrediting a digital forensic laboratory shows the world that your organization has an effective plan to drive out error and stretch towards continuous improvement to meet and exceed client needs.

An effective quality system will help your organization build a reputation for excellence where you will stand above others in the eyes and minds of clients. My experience has been, when the tough case presents itself, the first stop is to my laboratory where continuous quality improvement for client services is truly "Job #1."

Summary

Quality is attained through meeting client's needs, minimizing errors, and driving out critical defects in your product. A digital forensic laboratory must adhere to principles of forensic science using repeatable, measurable, and validated processes. Quality assurance includes the planned and systematic processes defined in your quality management system to meet your quality goals. Quality control is the methodology used to monitor product quality and the quality management system to determine if the quality criteria are effective.

Implementing a quality management system in a digital forensic laboratory is a process of establishing a cycle for continuous improvement to laboratory operations and meeting or exceeding client's needs. The quality system must be grown organically in the organization and must have full management support to be successful. Effectively launching a quality management system will require planning throughout the organization to obtain buy-in from management and employees.

One method for implementing a continuous quality improvement cycle is using the Shewhart-Deming model for Plan-Do-Check/Study-Act (PDCA) which is a continuous process for improving quality systems and product output. This cycle never ends.

Frequently Asked Questions

Q: What is quality?

A: Quality can be defined as meeting or exceeding your client's needs.

Q: What is continuous quality improvement?

A: A process that is driven by a cycle of planning, doing, checking, and acting (adapting) toward attaining a higher level of performance or product.

Q: What is the cost of implementing a quality System?

A: Properly implemented, a quality system is woven into the workflow of your business. It is an investment in your product quality. While there are overhead costs associated with maintaining the process and the review cycles, the cost of noncompliance of the product may have significant impact upon the business. Quality deficiencies at the Houston Police Department Laboratory and FBI Laboratory are examples of noncompliance.

Q: How long will it take to implement a quality system that derives benefits?

A: Benefits from planning a quality system are immediate. An analysis of your business processes to weave in a quality system that involves all layers of your organization will create a working atmosphere of trust and ownership. Ownership established early in the cycle of establishing a system for continuous quality improvement is in and of itself a benefit of the quality system.

Q: How quickly can a quality system be implemented?

A: It depends upon your organization; however, a quality system should be implemented incrementally to introduce the concept of quality processes and to avoid organizational shock to widespread implementation of new or refined processes.

Note

1. Shewhart, Walter and W. Edwards Deming. ASQ Handbook. American Society for Quality, 1999, 13–14.

Chapter 5

Balancing E-discovery Challenges with Legal and IT Requirements

Solutions in this chapter:

- **Drivers of E-discovery Engineering**
- **Locations, Forms and Preservation of Enterprise Electronically Stored Information**
- **Legal and IT Team Considerations for Electronic Discovery**
- **Steps to Becoming Discovery Ready**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

The ultimate goal of electronic discovery (e-discovery) is to provide electronically stored information (ESI) to a requesting party, whether it's the government, opposing counsel, or a third-party entity. As a former information technology consultant, it was easy for me to transition into the computer forensics industry. The progression from one to another allowed me to better prepare for a career in managing large and complex cases involving electronic discovery. One of the key elements of becoming a reliable source in the electronic industry is going to be your ability to translate technical information to ensure that you understand enterprise content management, data preservation, data collection techniques, culling, document review and productions. The objective of this chapter is to provide you with additional knowledge as to how to navigate a corporate infrastructure, and enterprise content, and address its links to e-discovery.

Drivers of E-discovery Engineering

"Discovery engineering" is a term that has been used in other fields for some time, but in our field, e-discovery engineering is a term that should be applied to individuals overseeing the entire electronic discovery process. E-discovery engineers are individuals whose primary focus is to alleviate some of the stresses on legal teams by understanding the technical intricacies of enterprise content, computer forensics, e-discovery processing, document review and productions. E-discovery engineering takes an individual with knowledge in the areas of computer forensics, information technology, records and information management, information lifecycle management, litigation support, and applies it to the sole purpose of plotting a course for the entire e-discovery process. Here is an explanation of what led to the coining of this phrase for this growing industry.

Tools & Traps…

Data Cataloging

Data cataloging is a necessary evil in the e-discovery industry these days. Locating ESI can be an overwhelming task for IT departments to accomplish. Through a suite of tools it is possible to locate ESI rather seamlessly; however, it is sometimes an expensive undertaking to manage these tools in-house. As professionals in the industry, teams of individuals are deployed to interview key players and run tools to assist this daunting

Continued

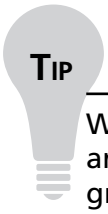
task. One of the products that could be utilized is DeepDive's DD300™, which allows consulting teams to locate and organize ESI in easy to understand reports. The product provides us the ability to index and search for potentially responsive information or allows us to extract useful reports that are used to communicate to the legal teams.

Storage

Ten years ago, the discovery industry was just that, the discovery of paper with an occasional hard drive requiring preservation and analysis. Today, paper is no longer the focus of this industry. The first factor affecting this growing industry is the advancement of the data storage industry. In the past we addressed storage devices with megabytes and gigabytes; but now storage capacity have greatly increased. It is often rare to address anything less than 100 GB and even more common to address multiple terabytes.

These terabytes result in millions of pages of documents, spreadsheets, presentations and many other types of files. However, the issue really relates to what do you preserve and how do you locate it. Through new tools, we now have capabilities to search for logical files on hard drives, which previously required us to forensically acquire a bit-for-bit image of a hard drive to preserve its contents. While there have been variations to allow us to collect logical files over the years, we were always required to sit and perform that work with the actual hard drive. Today, advances in our field's technology allow us to remotely collect this information.

Why is this technology so important to the e-discovery industry? It's simple. We do not need to access individual hard drives from a desktop or laptop every time we are required to collect ESI. Corporations are not preserving as much information overall, which is positive as it relates to their need to store that information as well as their need to manage it once its stored. And corporations could actually reduce costs through the use of these tools. For your clients or your company, cost is always the bottom line. So, consider your ability now to only collect, store, and manage what you need. While these mechanisms might not be full proof, they are a start to reducing the amount of information that needs to be processed through electronic discovery vendors. Therefore, while storage devices are increasing, the industry is shifting from the need to always physically acquire a hard drive to only collect potentially relevant information.



TIP

When identifying the locations of ESI, remember to always inventory on-site and off-site records and information. It is critical to the success of your program and, if you need to testify as to the steps you've previously taken, you want to be sure that you are up to date and aware of new system implementations, vendor changes, and retired hardware. And remember, the IT team is made up of your peers and you're only there to help.

Federal Rules of Civil Procedure

The second factor resulting in the need for advancements within electronic discovery is the changes to the Federal Rules of Civil Procedure (FRCP). Because of these changes, corporate and outside legal teams are asked about the complex details of corporate infrastructures, an area that is usually less than familiar to them. As a professional in this industry, your understanding of corporate infrastructures is essential and extremely beneficial to the entire legal team. The more the legal professionals understand, the better prepared they are during meet-and-confer meetings, negotiations, as well as being provided considerations from the courts. If counsel is able to present a strong argument about the complexities of the infrastructure as it relates to the production of particular types of records and information (RIM), it could be possible that less data is processed, which always results in a cost savings. On the other hand, legal teams with less of an understanding are at a serious disadvantage and this lack of understanding could significantly increase costs, time, and exposure.

Purpose

Some of the goals that have been attempted through the changes to the Federal Rules of Civil Procedure include cost containment, increased efficiency, and improved discovery management. Since I'm not an attorney, I look and speak about these changes only from a technical perspective. However, it is clear that the more an e-discovery professional understands about these changes, the better it is for the legal team. Some of the changes that might interest you are located in the following rules:

- Rule 16: Pretrial Conferences, Scheduling, and Management
- Rule 26(a): Duty to Disclose ESI
- Rule 26(b)(2): General Provisions Governing Discovery; Duty of Disclosure; Discovery Scope and Limits; Inaccessible Data
- Rule 26(b)(5)(B): General Provisions Governing Discovery; Duty of Disclosure; Discovery Scope and Limits; Claims of Privilege or Protection of Trial Preparation Materials; Information Produced
- Rule 26(f)(3) and (4): General Provisions Governing Discovery; Duty of Disclosure; Conference of Parties; Planning for Discovery
- Rule 33: Business Records as Answer to Interrogatories
- Rule 34(a) and (b): Data Archiving Requirements
- Rule 37(f): Data Resource Management/Safe Harbor Provision
- Rule 45: Conforms Subpoenas for ESI to other Rules
- Form 35: Report to the Court

So, what does it all mean? From a technical perspective, your ability to assist your legal team by answering questions as it relates to the production of ESI is critical. Legal participants typically understand the differences in producing ESI in native, TIF, or PDF formats. However, they might not understand what types of files can actually be produced and viewable in native format. In those cases, it is critical to have a team of individuals who can differentiate between an IPD file (Blackberry backup file) and a DOC file (Word document) to allow them to understand that opposing counsel might not be able to review an IPD file. While this might not be the most realistic scenarios regarding file types, it is of course, adding to your value as one of the technical professionals working on the case. The bottom line is, help the legal team understand as much as you can as it relates to technology. This will certainly prove useful in the long run and will save your company or clients money.

Costs

The third factor driving this industry's need for oversight is the cost-cutting measures required by corporations. While it might be easier to produce all file types to the requesting party, it is not a cost-saving measure for the client. Therefore, e-discovery engineers recognize that changing technology and e-discovery needs require thought-changing processes. Throughout my exposure to this industry, the increases in the amounts of data that are required have significantly increased. With new rules and as the courts understand this industry more, corporations are demanding cost-saving measures to alleviate their financial burdens as they relate to litigation requests. However, those cost-saving measures might not necessarily come from outside counsel due to a couple of reasons.

Although many attorneys have adopted a thorough understanding of the technical details as it relates to e-discovery, there are more appropriate individuals that should accomplish the tasks of navigating the infrastructure. E-discovery professionals should have the capability to structure a system that conveys as much technical detail to an attorney as possible and in a systematic format. Some of the information gathering phases include items such as the following:

- Custodian names
- Size of hard drives by computer
- Sample set of data from a statistically significant user population to identify averages
- Recommended method for collecting data from hard drives (for example, network forensic tools or local forensic acquisitions)
- Network share or home directory folder size allotments per custodian
- Potentially relevant network shares (by searching or filtering by custodian)

- File types of interest
- Costs to process file types of interest
- Identify reasonably accessible and inaccessible ESI (through on-site and off-site inventories)
- Restoration of backup tapes and costs (work with internal IT to determine best course of action)
- E-discovery processing methodologies that could be utilized

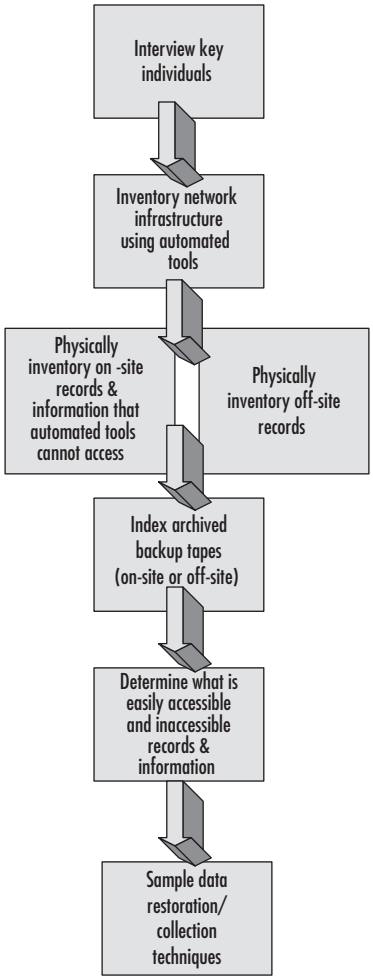
Shortly, we'll discuss how it is best to balance e-discovery challenges through enterprise content management, records and information management, information life cycle management, and technology and e-discovery tools.

Locations, Forms and Preservation of Electronically Stored Information

As mentioned earlier, locating electronically stored information can be a daunting task for IT departments. For example, take a corporation with 100,000 employees and consider how much information each employee creates and modifies on a daily basis. As IT departments are extremely savvy at managing their corporate infrastructures to meet daily business needs, they might not have the resources, expertise, or man-power to manage discovery requests. Therefore, one of the first steps you could take would be to inventory your ESI. However, it is difficult to find the time and the manpower when you already work sixty plus hours a week. If you or your client is in that situation, it is extremely beneficial to consider the deployment of tools, techniques, and people that might focus on those types of tasks. Many of my clients ask for an executive level overview of their infrastructure in addition to more detailed data cataloging. We'll talk more about data cataloging a little later in this chapter; but throughout the chapter I mention several tips, tools, and traps as it relates to this topic.

In Figure 5.1 a brief flow chart provides insight regarding the steps that are common in your initial steps of identifying the sources of ESI. While this is not an exhaustive chart, it provides a glimpse of the initial steps that are required to prepare for e-discovery. Later in this chapter we'll discuss interviewing, inventory, indexing backup tapes, and defining accessible versus inaccessible ESI.

Figure 5.1 Key Steps to Locating Electronically Stored Information


Interview key individuals
Inventory network infrastructure using automated tools
Physically inventory on -site records & information that automated tools cannot access
Physically inventory off-site records
Index archived backup tapes (on-site or off-site)
Determine what is easily accessible and inaccessible records & information
Sample data restoration/ collection techniques

Locations of ESI

Although ESI is available in many locations throughout an infrastructure, it has one common quality—it's electronic and is, therefore, stored somewhere. From a general perspective, ESI could be located on hard drives, floppy disks, CD-ROMs, DVD-ROMs, USB drives, backup tapes, optical storage devices, personal digital assistants, or mobile devices. Therefore, we have a good sense as to what we're initially searching for within an organization. But, we then have the issue of where all of these items are located and what is stored on each of them. At this time, it seems relevant to mention that enterprise content management (ECM) allows technical professionals to identify the locations of ESI through the use of e-mail archiving,

records and information management, information lifecycle, litigation hold, and document management systems. However ECM is not an extremely common practice in many corporations as it relates to e-discovery, so our teams typically need to be prepared to cover the ECM efforts through additional tools and techniques.

Forms of ESI

Forms of ESI really equate to the file types, which include the file headers and file extensions. The file headers and extensions distinguish between different file types. As other chapters will discuss, it is necessary to ensure that a file header is appropriate for the file extension. If a file extension has been altered, then the match no longer exists and a forensics team should be able to identify what its actual format is. However, in lieu of needing that type of integral service, it is critical to understand two other aspects of forms of ESI: (1) types of files that can be processed, and (2) potential metadata that is available within each type of file. While it's impossible to provide a comprehensive list of files or metadata, certain highlights do exist and should be considered.

File Types

As we discussed earlier, advancements in the storage industry have caused a surge in the need to collect terabytes of information, which equates to millions of files and even more pages to review. Discovery requests often request "all documents" or "all electronically stored information". As a word of caution, it is best to assist counsel by asking for the actual requests in addition to the complaint. With the knowledge you gain as a result of reading those documents, you are able to afford counsel the benefits of understanding what do they mean by "all documents" or "all electronically stored information". In many cases, you will be able to assist counsel with reducing the file types to more a commonly used list instead of "all". Common file types that might interest you could include documents, spreadsheets, presentations, portable document formats, e-mail containers and messages, text files, databases, Web pages, personal digital assistant backup and log files, instant messaging logs, server and workstation event logs, and Web mail.

Certain considerations should be given when determining your ability to extract Web mail, instant messaging logs, and event logs. First, it is critical to understand why Web mail is being requested. While we don't commonly see Web mail in over encompassing discovery requests, it is occasionally of interest. In these cases, ensure that your team is capable of restoring and converting the appropriate files. Web mail often comes in the hypertext markup language; but often times a proprietary file type is associated with it. In those cases, then it is necessary to ensure that you can appropriately handle this type of situation for the purposes of review.

Instant messaging is stored in logs and if the logging feature is not turned on, you'll only be able to review these files from a local hard drive image. However, if the logging feature is

turned on, then your electronic discovery processing team should be able to extract the data to post to a review site. In those cases, work with the internal technology team to restore this information to the appropriate servers.

Tools & Traps…

There's Always More Information

I know when we work with a corporation or a law firm that is attempting to preserve and collect ESI, we ask many questions. Because the IT department is often focused on managing the network availability, troubleshooting, or managing e-mail, they lack the time to properly identify sources of ESI. While it is common that asset management and inventory tracking systems allow you to track some of the assets/information within an organization, it is also common that this information is outdated. In these situations where an organization's IT department is unable to provide you detailed information as it relates to their locations of ESI, it is critical to interview as many technical individuals as possible and to work with them to locate as many forms of ESI as possible.

Metadata Fields

As you know metadata is "data about data". What makes data about data so important then? It provides useful details as it relates to the creation date, modified date, accessed date, creator names, and sent and received dates, in addition to other facts. Some of the fields that are typically seen within the metadata are different for electronic documents as compared to electronic mail. Table 5.1 provides additional information about metadata.

Table 5.1 A Brief List of Common Metadata Fields

Field Name	Description
File Name and Path	The full path and file name of the file.
File Name	The name of the file.
File Path	The full path of the file.
MD5 or SHA1	The unique identifier of the file.

Continued

Table 5.1 Continued

Field Name	Description
Document Type	The file extension that was detected during processing, regardless of the file header.
Email	Provides details as to whether the file is an electronic document or an electronic message.
Parent	This determines if the file has an attachment.
Child	This field distinguishes if the file is an attachment.
Duplicate	Determines if this file is a duplicate among the entire population files or a subset of the population.
Date Accessed	Last access date of the file.
Date Modified	Last time the file was written to
Date Created	The date in which the file was created.
Text	The extracted text from the file.
Author	The creator of the file.
Character Count	Number of characters in the file.

Additional metadata that is available from the file could be comments or dates from the metadata itself. It is common for additional information to be available when processing foreign languages. This is often times due to the need to process a Unicode-based character set and customized processing steps could be required. Other metadata that is extracted could be the document's page count, revision number, subject from a message, titles, word count, company name, and document manager's name. The processing industry allows certain levels of customization when extracting the metadata from an electronic document or an e-mail, but there are certain limitations. It is critical to understand the limitations from each service provider to make informed decisions about the processing of electronic files.

Legal and IT Team Considerations for Electronic Discovery

It is becoming more common for corporate legal teams to dedicate IT professionals to their groups to assist with litigation hold efforts in addition to preservation and collection duties. However, it's not a standard exercise at this time either due to budget constraints, a lack of understanding of the value of IT professionals within legal departments, or corporate

structure. A good intermediate solution is to form an e-discovery team within your corporation or to hire a firm that can assist you with forming an e-discovery team. By taking those steps, your team will better prepared for the daily discovery request challenges and will provide you a more defensible and admissible game plan over time.

IT Members within the Legal Team

Let's take a look at the role of an IT professional within a corporate legal department. In a perfect world, this individual should have the enterprise content, data preservation and collection, electronic discovery, vendor management, legal and training experience required for your group to benefit from this individual. Typically, consultants with several years of experience with a combination of these skills make excellent legal team IT support in response to e-discovery requests.

The concept of an in-house e-discovery team is growing as the industry continues to flourish. Therefore, if you plan to integrate an e-discovery team to your efforts, it is recommended that you consider multiple aspects of the team. First, build a team with at least one executive to make sure your efforts are supported at a high level. Second, ensure that representation from multiple divisions or business units are represented. Don't leave out human resources because they will be critical to identifying terminations and could be very useful for communication purposes. Next, build a team with at least two technology professionals, one to two legal staff members, the records and information manager, enterprise content managers, and business unit leaders. Obviously, the goal is to create a team that will contribute to the success of your program. Fourth, identify a team leader to ensure that the goals are met for this internal organization. Without setting goals and adhering to them, you're better off not creating a team at all. And finally, kick off the meeting with an agenda, projected implementation timelines, as well as an overall strategy. This will make for an organized kick-off and a successful team initiation.

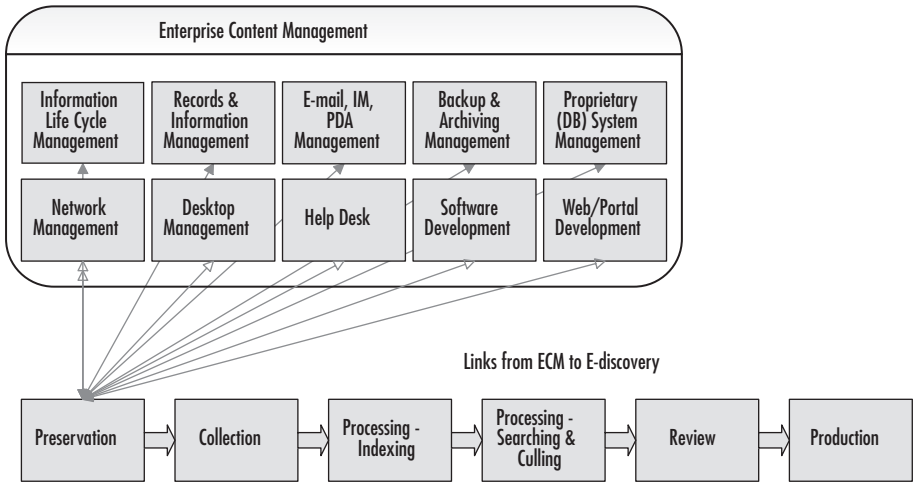


TIP

When selecting a member of the IT team to join the legal team in an effort to support your discovery requests, choose an individual with a number of years in the technology field designing, administering or supporting corporate infrastructures. Ensure that this individual understands data preservation, collection, culling, and review. This will be essential to selecting the correct person.

Over the years while managing IT teams, investigating computer cases, or responding to discovery requests, I've seen a number of different scenarios within corporations. Some of them possess a high level of sophistication that allows us to do our job efficiently, while others need some help. Below is an illustration that depicts some of the aspects of enterprise content management and where their link to the e-discovery process begins. While team members should have involvement throughout the process, it is critical to have a starting point. It might surprise you that the starting point is when the litigation hold has been received. Without the ECM team's involvement, adhering to a litigation hold is difficult and sometimes impossible. Figure 5.2 provides a glimpse of when it should be necessary to involve team members that manage an enterprise's content. However, the illustration only depicts parts of the folks managing the enterprise content and is not meant to be exhaustive.

Figure 5.2 The Point of Initiation: Data Preservation Is the Key



Managing the corporate infrastructure is a key aspect of the enterprise content team; however, if you start to involve them in assisting you with your discovery requests, certain individuals will be crucial to the process.

Records and Information Managers

Corporate technology teams increasingly have a need to integrate records and information managers into their programs as electronic discovery requests increase. Records and information managers (RIM) traditionally worked with paper records and information, but were not involved in the management of ESI. Therefore, despite RIM being a widely recognized profession, technical and legal teams are just starting to understand their importance amongst their team. The utilization of a RIM benefits all parties involved since these individuals understand the need to organization records and information, but also might have additional insight into

their locations, size, and purpose. Again, that's from a hard copy perspective. From an electronic perspective, it is always good to utilize an individual with a strong technical background in addition to the records and information management knowledge base. Let's discuss in greater detail the value a records and information manager could have for you or your client.

Logistically speaking, the records and information manager could assist by establishing a records and information program for hard copy and electronic copy. One of my clients once called and said we need to do a project. She proceeded by stating that she had an outdated records and information management system along with an outdated manager. While I normally hadn't received these types of calls in the past, the shift has gone from one of a business function to manage records and information to that of a legal. She continued to say that 80 percent of the hard copy records were being scanned and imaged into an enterprise content management system. The other 20 percent were simply being stored in individual boxes. They had not been scanned into any system and had not been inventoried. This led my team to believe that the reason for the call was much more about business process improvement than the functionality of the ECM system in place at the time.

The moral to the story above is that your goals within records and information management should be tied as much to managing the discovery process as it is to the business process. We continued on with this client and found that the ECM system that was implemented had such a larger amount of functionality than originally anticipated. Not only was the client able to adjust their expectations regarding records and information management, but they were able to utilize this system to assist in future e-discovery.

Information Lifecycle Managers

I often have a tough time keeping the positions straight within the corporation. Similar to the records and information managers, the information lifecycle managers (ILM) are also critical to the success of your electronic discovery program. Without individuals and their knowledge of storage devices and mechanisms, the management of your storage mediums would be nowhere. However, information lifecycle managers are not greatly utilized during the e-discovery process. Therefore, we need to change something. These individuals are extremely beneficial as you acquire a need to hold more and more ESI. Their knowledge of storage management and their abilities could actually allow you or your client to hold data on a specific area of your network. One of my colleagues recently coined the phrase, "Hold Area Network", which addresses a network storage location that allows a corporation to store ESI due to a legal hold. While the term itself is self explanatory, the steps required to create a hold area network might not be as apparent.

E-mail, IM, and PDA Managers

When handling discovery requests for my clients, one of the easiest parts of my job is requesting e-mail from an e-mail administrator. That is, if their e-mail archiving solution is

not stored offsite and provides no direct access to the actual e-mail server any longer. I have a couple of scenarios to share with you to make your preservation and collection efforts a little easier as it relates to e-mail collections.

Some time ago I had a client who called to ask for assistance with responding to a litigation request. This request specifically asked for all communications for a particular business unit from June 2004 through June 2005. While it was possible for our client to comply with the request and provide all of these communications, it was found to be somewhat burdensome. The infrastructure implemented by the IT department had multiple issues with it. Because the organization was so large, the e-mail archive system that was implemented in 2003 stored the content off site, and there was no direct access to the off-site e-mail servers. Since the e-mail servers were no longer actually storing e-mail content, they simply served as the shell of the e-mail infrastructure. Although our client was relatively large (20,000 plus employees) in nature, they did not have the manpower to address the litigation request. Thus, our team enters the picture to address the issue.

Upon entry to this client, we were informed that there were 30 custodians that were relevant to this request, e-mail messages were saved for each custodian in the allotted time frame, and the e-mail archiving solution caught inbound and outbound messages prior to the custodian receiving it. What we didn't know at the time was that the IT department had not updated the archiving solution and that extractions of e-mail messages had actually encountered errors for several months, but that the internal resource had not reported the errors. Therefore, virtually every production within an unknown period of time prior to our arrival could have been incomplete. The point of the story is, if you are responsible for restoring information to provide to a requesting party, make sure the applications you're using to restore that information is working properly.

This brings up another point regarding the restoration of information. One of the most neglected areas when responding to discovery requests is the use of tracking. If you restore it, track it. Not only does this provide you with an inventory of the information you are providing, but it also allows you to account for the original information linked to some form outside of the original storage area. Tracking systems can be as simple as the samples shown in Table 5.2.

Table 5.2 A Sample Tracking System to Highlight Gaps

Source Name	Message Count (Source)	Message Count (Restored)	Delta
Custodian1.pst	262	185	(77)
Custodian2.pst	269	269	0
Custodian3.pst	698	667	(31)
Custodian4.pst	401	401	0
Custodian5.pst	397	394	(3)

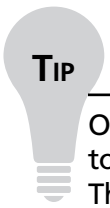
Simple gap analysis spreadsheets afford a couple of vital functions when restoring information. First, the above spreadsheet determines that we did not extract the same number of e-mail messages that were noted within the archiving system. While it's possible that you may never be able to restore the missing messages, it is critical to attempt the restoration process again, for example, with the custodians specified above. Upon a second attempt of extracting the messages from the system and if unsuccessful, make a comment regarding the errors or reasons for the failures. This provides detail to the legal team that they might require when negotiating with opposing counsel or discussing with the courts. The more detail you have, the better prepared your team will be.

Second, tracking is an excellent tool when you restore information and have a need to provide it to a third party, whether an EDD processing service provider, document hosting provider, or opposing counsel. The value is simple here. If you have a tracking mechanism and you are able to provide the outside parties with accurate information, questions will not arise regarding the information you provided. If you are not prepared and don't have tracking mechanisms in place, then you run the risk of scrutiny, additional costs to find the gaps, as well as lost time and a need for additional resources. All of those factors make this a necessary and critical step in the e-discovery process.

Backup and Archiving Managers

While you might have an e-mail archiving system that requires management, I refer to backup and archiving separately from the e-mail infrastructure. Although there are many overlapping components, we'll simply consider two aspects in this section: (1) backup management as it relates to backup tapes and optical storage mediums, and (2) off-site storage management.

Nearly two to three times a month my company receives a call from clients saying they have backup tapes and they don't know what's on them. For many reasons, IT teams lose track of the contents of backup tapes. In some cases its just poor tracking systems; in others it was just easier to send these tapes off site because they didn't have the resources to inventory them; possibly you did not work for the company at that time and your predecessor didn't inventory anything, or may be some kind of disaster caused the loss of the inventories of those tapes. Regardless of the reason for the lack of inventory of backup tapes, it simply has to be done and it has to be done now. Newer technologies allow you to index and inventory tapes in a fast and simple fashion.



TIP

Off-line tape discovery is essential to identifying locations of ESI. One of the tools that prove extremely useful in this process is a product by Index Engines. Their indexing process for e-discovery allows you to directly index tapes from multiple tape manufacturers and backup software. The product is easy to deploy as well as easy to use. Through the use of products such as this, you will be able to reduce the expenses that were once needed to simply index an offline tape.

Are You Litigation Ready?

Becoming litigation ready is no small task. However, you need to start somewhere. The starting point is the tough part for corporations that have been responding to discovery requests for some time without having a sound process in place. I understand you're probably saying, why wouldn't you have a sound process? Well, I can tell you from my experience, discovery requests come upon a corporation so quickly or an internal investigation begins and it is very difficult at that point to back up to consider the necessary steps to really be litigation ready. So, let's look at two different scenarios: (1) steps you can take if you are served with a request and don't have a process in place, and (2) steps to prepare.

Served with a Request

It is always easy to say, why aren't you ready for this request? But, so many companies really aren't ready. You receive a request on a Monday morning (typically they have a habit of showing up on a Friday afternoon), assemble your legal team, and look at each other. Or, outside counsel calls and states that you need to provide ESI for 100 users, including e-mail messages from a previous time frame, active hard drives need to be preserved, and backup tapes are called upon in the request. You start to think, where do we begin? With either of those scenarios the planning phase goes out the window, but the following guidelines could assist you.

Contact Your Chief Information Officer or Equivalent

You really need to have a trusted source who understands the key players that are required to collect this information. If your internal teams are completely tasked or have not dealt with a request like this before, contact a professional.

Be Prepared to Field Questions from the Professionals

After you contact the professionals, they are likely to ask the following questions:

- What form would the opposing party like the ESI (for example, native, TIF, or PDF with or without load files)?
- What does the complaint involve (typically the nature of the suit will assist in dictating some of the necessary steps)?
- How many custodians are involved?
- Are you being asked to forensically acquire desktop/laptop hard drives?
- If yes, then who will perform this work? This topic is covered in other areas of this book; so we won't elaborate on responses or technicalities surrounding the forensic acquisitions.
- Do all of the custodians still work for the company?

- Are your IT efforts centralized or decentralized?
- When is the request due or what is the proposed date?
- Do you have an IT team member that could assist in the efforts to collect this information? This is critical if e-mail, instant messaging, Bloomberg Mail, and others are due with the request.
- When would you like to get started?

Be Prepared to Ask Questions

As the client, some of the things you should ask include the following questions:

- Could you provide us with an overview of your services?
- Could you provide an estimate to complete the initial phase of the work?
- Do you have a standard chain of custody (CoC) form?
- How long have you been performing this type of work and what levels of expertise do you have?
- Can you provide us with references? Although this part is sometimes difficult since you want to get started right away, the outside company should have some credentials that acknowledge the companies accomplishments, publications, speaking events, and so on. References are a good indicator that the company is a thought leader in the industry.
- Can you provide us an approach? This allows you to see the flow of the work that needs to be completed.
- What is the level of credentials of the individuals that will perform the work? Can you provide resumes for your key individuals?
- Do you perform a background check on your employees?
- When can you get started?
- Who do you need access to within the company to obtain the information required?

Although the preceding guidelines are not intended to be comprehensive, they are a good starting point. So, what happens if you actually get started? Following are some steps that could provide you some oversight as to how the process flows:

Interviews

Here are some topics that you need to discuss with technology team members:

- Custodian locations
- Standard desktop/laptop hardware

- Standard desktop operating systems
- Potential locations of ESI (hard drives, servers, off site)
- Average hard drive sizes
- Shared network and home directory locations
- Web portals
- Instant messaging details
- E-mail
- Standard desktop applications

You also need to discuss these topics with your legal team:

- Potential scenarios for collecting ESI
- Potential pitfalls and risks
- The next steps

With key custodians, you'll need to discuss these topics:

- Data storage locations
- Number of desktops/laptops
- Special/proprietary applications
- Purchasing process for technology
- Number of computers used during the time frame in question
- Level of computer knowledge and usage

Inventory

When taking inventory at on-site and active locations, you should obtain a collection of inventory records about the ESI in question. If inventory records do not exist, take an inventory of potentially relevant locations of ESI.

When taking inventory at off-site and archived locations you should obtain a collection of inventory records about the ESI in question. If inventory records do not exist, take inventory of potentially relevant locations of ESI. If backup tapes are in question and an inventory/index does not exist, then index and inventory potentially relevant tapes. If optical storage platters exist, determine what ESI resides on the platters and inventory, if necessary.

When taking inventory of retired hardware, you should determine if retired equipment is potentially relevant based on history of use, preserve hard drives if an inventory is not available and costs are too high, and provide file listings through forensic methods if required.

It makes sense to inventory every system available to your corporation. Following the inventory phase of on-site and off-site records and information, identify standard operating systems, standard office productivity applications, proprietary systems, special applications assigned to certain divisions, executives, or individuals based on their job responsibilities. Next, it is often a good measure to follow up on your interviews following the identification of standard and proprietary information. At this point, you may have identified new individuals that require an interview.

One of the best methods in identifying ESI is to conduct interviews. These interviews, even if conducted by internal employees, allow the opportunity to note and identify items that might be normally overlooked. As a seasoned veteran of conducting these interviews, the best method to follow is to have more than one person answering the technical questions. This allows for multiple points of view as well as an elaboration. Typically an interview will take three to four hours initially with, possibly, 12 to 15 hours of follow up questions as well as the collection of documentation. This timeline typically depends on the complexity of the company's infrastructure, the knowledge of the IT department, as well as the organization of ESI.

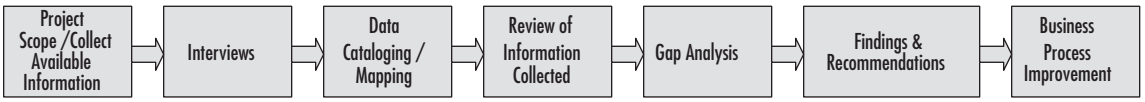
NOTE

While conducting interviews with key individuals, it is critical to take lots of notes. In addition, the interview process should be conducted by two professionals to ensure that missed details by one person are not missed by the other. This allows somewhat of a stopgap to document particular details and information that could be critical at another stage of the case.

Discovery Readiness Planning

If afforded the opportunity to plan for discovery requests, take it. This is too often a missed opportunity because litigation persists and time is no longer available to plan. Figure 5.3 illustrates a high level of overview of the steps performed during the litigation readiness project. We'll review this in a little more detail in the following pages of this chapter. Take note that some of this information is very similar or the same as the "Served with a Request" section.

Figure 5.3 Critical Steps in Becoming Discovery Ready



Project Scope/Collect Available Information

Project scoping is one of the oldest steps in project management, but it is critical for these types of projects. It further defines key players along with expectations. For example, if you have a client or employee that wants to plan for litigation and they ask you to index every record on the network, then the expectations regarding time might be different than if someone requests a folder that lists a particular server farm. The slight nuisances of what your team and the tools available can do in a cost effective manner are critical and you need to understand how to manage them in a project like this. From a discovery perspective, these projects are short term and should be handled with the utmost sense of urgency to provide the deliverables.

Request documents from your company or client similar to the following:

- Network topology
- Litigation hold request communications/policies
- ESI retention and destruction policies
- Asset management reports
- On-site inventory records as they relate to hardware, software, ESI
- Off-site inventory records as they relate to archives and backup ESI
- Previous discovery effort reports
- Previous tracking mechanisms
- Previous forms of metadata available for review
- Previous collection, processing or review details available

Interviews

This interviewing phase is very similar to the steps listed above; however, this phase is also an opportunity to focus on a broader scope. These interviews should be exhaustive in nature and should ask details about the following areas:

- General information number of employees, locations, and so on)
- Policies and procedure information (termination, retirement of equipment, training for new users, data classification)

- File structure (home directories, shared drives, databases)
- Network access (locations, FTP, portals, outside user access)
- E-mail (number of e-mail servers, type of e-mail system)
- Logging capabilities (instant messaging, security)
- Instant messaging policies and management (enterprise, configurations, backups)
- Voice mail and telephony (Internet-based or PBX)
- Off-site and/or external data storage (remote employees, list servers)
- Inventory and asset management (storage locations, total number of desktops/laptops)
- Records and information management (policies, systems)
- Document management (policies, systems)
- E-mail archiving (policies, systems, storage locations)

During the interview phase, it is critical that you discuss your company's current processes for preservation, collection, processing, hosting, and production. Questions that should be asked include the following:

- How do you currently address a litigation hold? Discuss communications, processes, and policies.
- What type of chain of custody is utilized for your collection practices?
- What is the current method utilized for collecting hard drive ESI?
- What is the current practice used for collecting ESI from servers?
- How are you collecting ESI from archived and backup or off-line media?
- What is the current procedure used to transmit the collected ESI to a service provider for processing?
- Once the ESI is processed, where is it hosted for review?
- What type of hosting platform are you utilizing for the review?
- Once the ESI is processed and reviewed what is the typical form of production?
- How do you manage your cases in terms of ESI delivered to requesting parties and internal review statistics?

Data Cataloging/Mapping

For the technical folk, data cataloging and mapping is the fun part. During this phase, your team should create high-level and detailed network topologies and user file inventories. As you'll see below, there are tools like DeepDive DD300 that can assist you with creating user file inventories. The screen capture shown in Figure 5.4 provides you with a high level overview of locations on the network that have been indexed. It also identifies the number of documents that have been indexed, the ones that cannot be indexed, and the amount of ESI you'll encounter at each location.

Figure 5.4 DeepDive DD300 Indexing Report: File Locations for Data Cataloging

Deepdive - Indexing Report

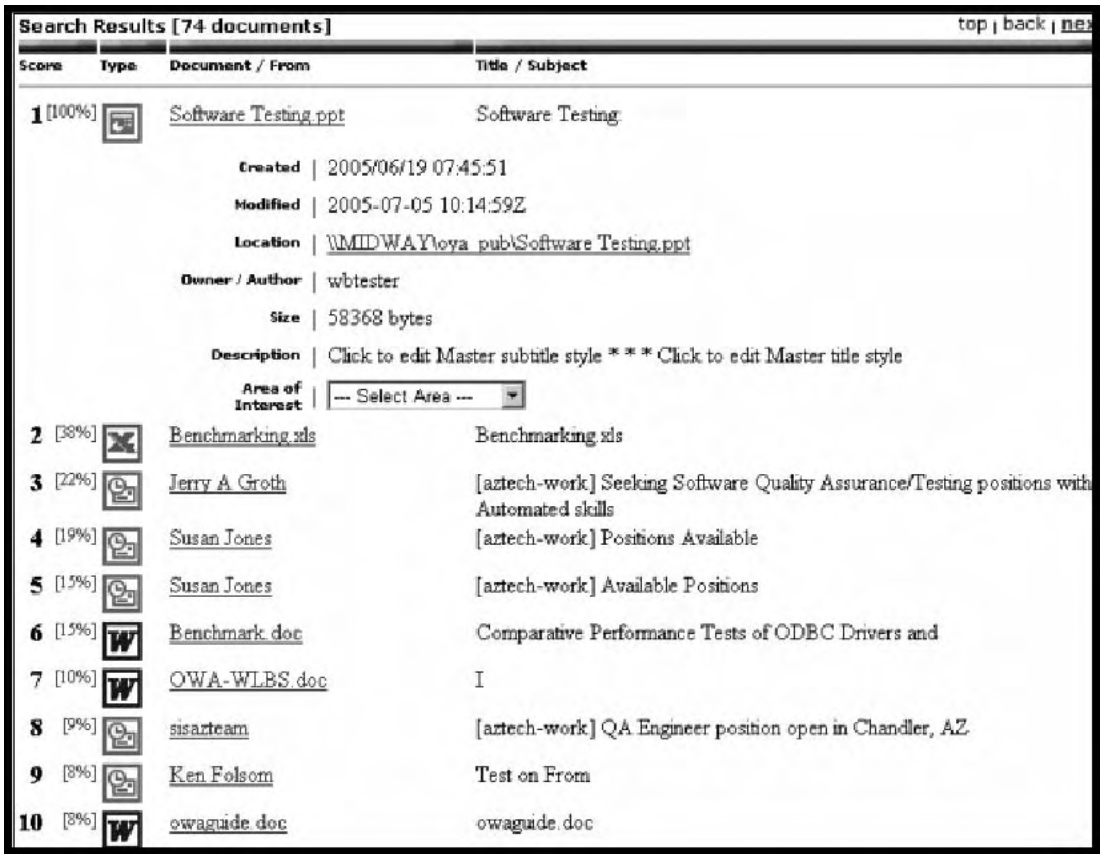
Summary Information

Report Date:	**Wed Apr 11 12:00:22 2007**
Indexing Vendor:	**dtSearch 7.30 (7301)**
Start of Indexing:	**Wed Apr 11 11:57:42 2007**
End of Indexing:	**Wed Apr 11 12:00:22 2007**
Elapsed Time:	**2 mins 40 secs**
Indexing Type:	**Incremental indexing run**
Maximum Documents:	**5000000**

Indexed Locations	Index Size (bytes)	Documents	Successes	Failures
\\OLRSAP25\C$	26414051	2420	2419	1
\\OLRSAP25\F$	377031	0	0	0
\\OLRSAP25\Q$	565914	4	4	0
\\OLRSAP25\R$	377031	0	0	0
\\OLRSAP25\S$	1523190	78	78	0
\\OLRSAP25\Share	557716	4	4	0
\\OLRSAP25\T$	377031	0	0	0
	30191964	2506	2505	1

Furthermore, additional reporting that might interest you includes metadata and actual file types. As seen in Figure 5.5, a common screen capture within DeepDive's DD300 allows you to capture information about file types and their associated metadata.

Figure 5.5 DeepDive File Listing and Associated Metadata Screen Captures



Another tool that is useful once you've identified the locations of ESI is ONSITE[3]'s e-discovery Estimator. This tool assists you in scoping the processing requirements of potential e-discovery projects and, consequently, the cost that may be involved in processing data for e-discovery. If you have already known the costs that your service provider charges to process certain types of ESI, then this tool is a great mechanism for identifying the volumes.

TIP

Data cataloging is a necessary evil in the e-discovery industry these days. Locating ESI can be an overwhelming task for IT Departments to accomplish. Through a suite of tools it is possible to locate ESI rather seamlessly but is sometimes an expensive undertaking to manage these tools in-house. As professionals in the industry, teams are deployed with a team of individuals that interview key players and run tools to assist this daunting task. One of the products that could be utilized is DeepDive's DD300™, which allows consulting teams to locate and organize ESI in easy to understand reports. The product provides us the ability to

index and search for potentially responsive information or allows us to extract useful reports that are used to communicate to the legal teams.

Review of Information Collected

Once you collect the information required to officially assess your company's e-discovery needs, you'll need to review it. The review process is fairly straightforward and flows directly into providing a gap analysis. The review should consist of the following:

- Interviews to determine additional questions
- Data cataloging/mapping to identify other areas of interest
- Methodologies and/or service providers preserving, collecting, processing and hosting ESI

The review of the information phase allows your team to identify additional information that could be required from your company or client. It assists you with moving into the gap analysis phase, where you will be better identify the gaps in the e-discovery process or new technologies available to assist your discovery readiness plan.

Gap Analysis

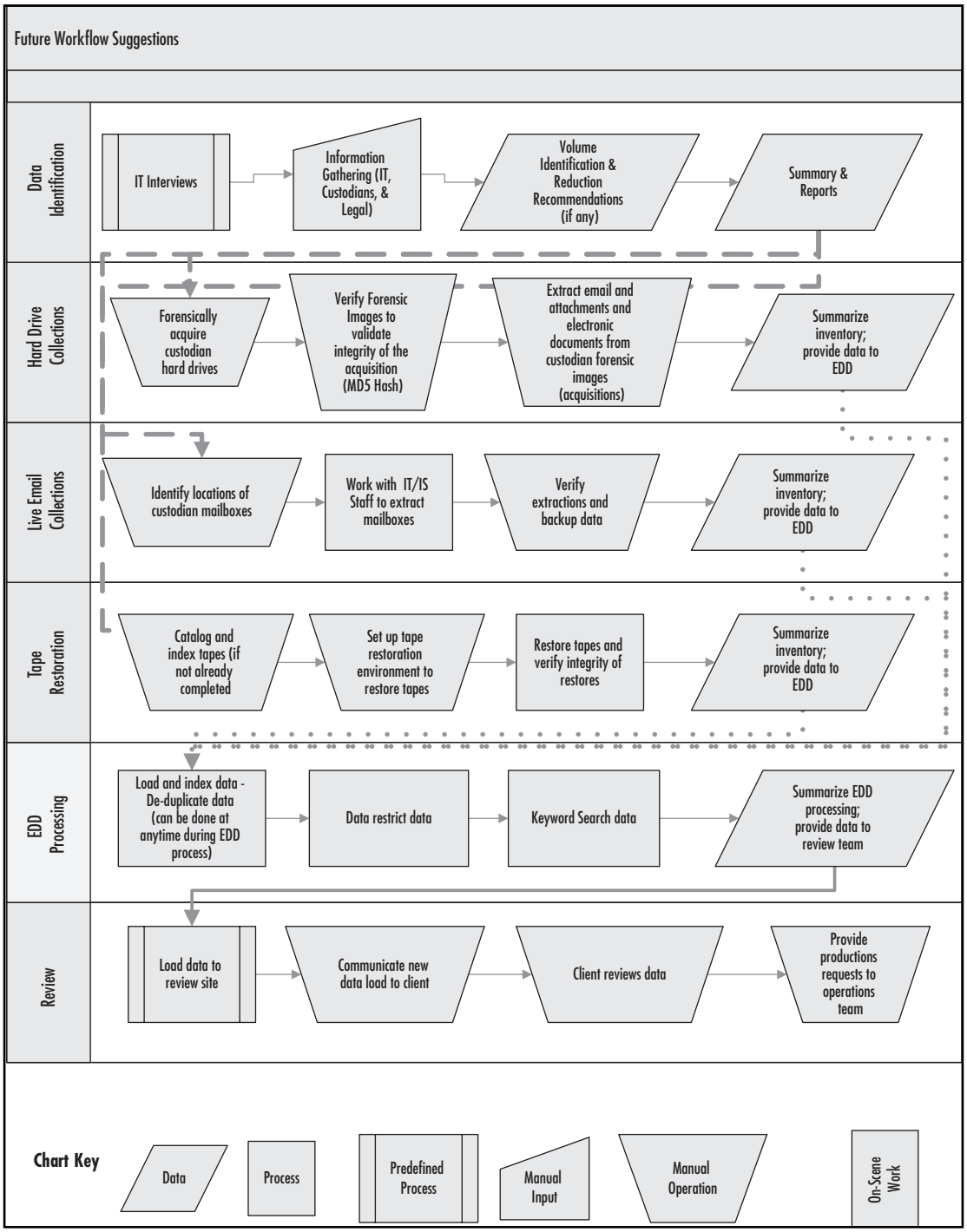
During the gap analysis phase, it is critical to note your current practices as they match up to recommended or best practices within the industry. This phase is essential to improving your practices. Occasionally counsel asks that we simply identify the best practices in the industry to avoid the issue of stating that there is a gap in the process. However, it is best to actually work through this phase since crucial information could be utilized to improve your practices.

During the gap analysis, you are able to further review the information collected during the preceding phases to capture details such as issues with your collection methodologies, collections that might not be complete, additional off-site or archived ESI that hasn't been addressed in the past, less than reliable practices, and other technologies that might be useful. This phase is completely customizable and should include a spreadsheet with comparisons of "where we are now" versus "where are we going". This simple, but useful step provides guidance into the findings and recommendations phase.

Findings and Recommendations

During the findings and recommendations phase, your team should be able to address the next steps. In many situations, my clients are able to tell me how they currently handle ESI and its production, but they might not have the experience to identify with better practices. Therefore, my goal is to lay out a flowchart that will assist them with creating a more cost-effective and streamlined e-discovery plan. In the screenshot shown in Figure 5.6, you'll see how I created a working flowchart to graphically display a new process. Pictures are very useful because we are able to display their current workflow versus a futuristic state. We then take that and point out the gaps from the previous phase. By doing this, we are able to create a plan to implement change to improve this process.

Figure 5.6 Future Workflow Suggestions: Data Identification, Collections, Processing and Review


Future Workflow Suggestions
Data Identification
IT Interviews
Information Gathering (IT, Custodians, & Legal)
Volume Identification & Reduction Recommendations (if any)
Summary & Reports
Hard Drive Collections
Forensically acquire custodian hard drives
Verify Forensic Images to validate integrity of the acquisition (MD5 Hash)
Extract email and attachments and electronic documents from custodian forensic images (acquisitions)
Summarize inventory; provide data to EDD
Live Email Collections
Identify locations of custodian mailboxes
Work with IT/IS Staff to extract mailboxes
Verify extractions and backup data
Summarize inventory; provide data to EDD
Tape Restoration
Catalog and index tapes (if not already completed
Set up tape restoration environment to restore tapes
Restore tapes and verify integrity of restores
Summarize inventory; provide data to EDD
EDD Processing
Load and index data - De-duplicate data (can be done at anytime during EDD process)
Data restrict data
Keyword Search data
Summarize EDD processing; provide data to review team
Review
Load data to review site
Communicate new data load to client
Client reviews data
Provide productions requests to operations team
Chart Key
Data
Process
Predefined Process
Manual Input
Manual Operation
On-Scene Work

Business Process Improvement

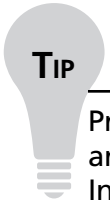
Upon the completion of the above steps and acceptance by senior management, the business process improvement (BPI) phase will set the plan in motion. Just as this step is critical in the success of other plans throughout the company, it is critical to the success of your e-discovery program. The BPI phase is easy to implement once you have a game plan; however, buy-in from senior management is also critical. Regardless of the changes you decide to make, the costs can be quite large. If you plan to implement an in-house e-discovery program, the costs will be much greater than turning to outside support. Therefore, it is necessary to identify costs throughout the process to ensure that your management team understands the implications.

E-discovery Tools

Now that you've taken the proactive measures necessary for a sound e-discovery program, its time to consider some of the tools you can utilize to process your data. While I work for a company that utilizes proprietary software to process client data, some companies utilize off-the-shelf software to accomplish the same task. When assessing off-the-shelf software to accomplish your e-discovery processing, ask the following:

- How long has your product been on the market?
- How do you handle service packs or upgrades?
- Does your software upgrade require a hardware upgrade?
- What type of training is involved if I utilize your software?
- How many users can concurrently utilize the software?
- How much does your software cost? Are there annual maintenance fees?
- How much hardware will I need to efficiently run your product?
- What is the deduplication process utilized by the product?
- What reporting is available as a result of the import and reduction of data sets? Can I receive a sample report?
- Do you have a near deduplication feature built into the product?
- Can I convert voice files to text for searching?
- What is the backbone of the product?

The above list is only a glimpse of the questions that should be asked. Consider all of the aspects that are critical to the success of your processing to ensure that you are more than comfortable with the product's capabilities. Also, ask if a sample set of ESI could be processed to provide additional comforts. Compare all of the products side by side to get a realistic comparison. While nothing is completely foolproof, these steps do implement a little more due diligence on your end.



TIP

Processing tools should effectively allow you to import ESI, track its status, and reduce the set of ESI through deduplication and other culling processes. In addition, the output of ESI should be a usable fashion to ensure that it can be transferred to a review tool or to the requesting party in an acceptable format. One tool that has been on the market and is well-known in the e-discovery processing world is called Workbench, a product by Attenex. This product allows you to seamlessly import data for processing, report on its status, and provides a number of output formats. It also further automates technical document processing needed to catalog, extract, suppress, index, analyz, and prepare electronic files for review.

Summary

As you can see from this chapter, implementing a discovery readiness program is not that straightforward. However, by utilizing trusted sources and tools, you are one step closer to implementing a program that works. As the industry continues to evolve, additional tools and techniques will be available on the market. In the meantime, verify that your processes and techniques are valid. And it never hurts to read current case law to check up on new findings.

As we venture further and further down the road in e-discovery, our methodologies, tools, and techniques will change. E-discovery engineers will also become more prevalent as the industry evolves with their new methodologies, inventions, and ideas. If you are just entering this industry and need a path to follow, becoming an e-discovery engineer could be fruitful. It's a changing industry and is in need of talented technical resources. By becoming an e-discovery engineer, you could move to the forefront of this industry.

Legal and technical teams need to bridge the gap by creating a process before venturing down the road of e-discovery. Unfortunately, the process is not perfect. There are preservation, collection, processing, and review issues, but through team effort, these issues can be minimized. Through team planning, you will be one step closer to implementing a streamlined e-discovery process. But, don't stop there. Remember to revisit the plan every three to six months since new technologies and methodologies will be available.

Understanding an enterprise's content is a great step to assessing the volumes of ESI and their associated locations. As we mentioned earlier, look for key entry points to utilize the IT staff. Work with the enterprise content managers to learn more about your records and information, e-mail, archiving, off-line storage, and other key areas. Ask questions, and shadow these individuals if necessary. The more you learn, the better prepared you'll be when needed. And, as you work through this information gathering phase, form useful relationships. These relationships will greatly assist you with the preservation and collection phases since the folks you'll need to call upon will understand your needs.

Discovery readiness will put you one step closer to being litigation ready. Through some of the techniques and tools identified in this chapter, you'll be able to identify the gaps in your process. Don't forget that interviewing is essential to learning about the infrastructure, and users will hold a great amount of the answers. If you simply interview the IT team, you might miss critical information that could lead to additional sources of ESI. While the IT teams typically hold the key to this castle, often users have other sources. Don't get caught in the trap of only having enterprise IT details. Keep your eye on the users and gather the information they are willing to share.

Frequently Asked Questions

Q: What is the best way to get started with planning for discovery requests?

A: If your company does not possess any formal documentation or inventory records, get started now. These are critical in understanding the volume of information that will be required for each litigation request.

Q: What components should I consider when looking for a data cataloging tool?

A: Searching technology is probably the key component, but is not as critical as the ability to search a network and hard drives without having to manually share out information. Like many of the network collection tools available on the market, it is critical for data cataloging tools to provide the ability to search a network without having to share out drives.

Q: When should I engage an enterprise content team member in the e-discovery process?

A: It is best to engage as many members from your ECM team at the beginning of the e-discovery process.

Q: What are some of the metadata fields that should be considered when processing ESI?

A: Key metadata fields that should be considered include, but are not limited to the unique identifier, hash value, file name, file path, parent-child relationship when considering e-mail, file size, created date/time, modified date/time, last accessed date/time, text extracted, and the author.

Q: What are the key steps to ensuring that you are discovery ready?

A: Scope the project and gather information, interview key and ancillary individuals, catalog the data on the network, review your information, run a gap analysis on the information collected, report on your findings and recommendations, and never forget to implement the business process improvement steps. The key aspect to discovery readiness is using the information you collect to improve those processes.

Chapter 6

Forensic Software and Hardware

Solutions in this chapter:

- **Forensic Software Tools**
- **Forensic Hardware Tools**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

In this chapter we introduce to the hardware and software-based forensic tools. Due to the large number of hardware and software tools, few can be covered in great detail.

The second part of this chapter introduces the physical hardware tools available to aid the forensic practitioner. Tables are included within the chapter that supply a brief description of the salient features of each tool.

Part 1: Forensic Software Tools

This section summarizes the features and advantages of a large number of software forensics tools. For detailed information and technical reports it is always best to view the vendor Web sites as well as organizations that conduct technical reviews and evaluations such as National Institute of Standards and Technology (NIST). The Computer Forensic Tools Testing project (CFTT) web site contains additional valuable information:

- http://www.cftt.nist.gov/disk_imaging.htm
- http://www.cftt.nist.gov/presentations.htm
- http://www.cftt.nist.gov/software_write_block.htm

The information presented in this chapter is heavily based on the assertions of the various vendors who make the products listed in the chapter. Much of the information has been taken from the vendors product sheets. The Computer Forensic Tools Testing project is a good source of comparative data when deciding between these vendors.

Visual TimeAnalyzer

Visual TimeAnalyzer, shown in Figure 6.1, automatically tracks all computer usage and presents detailed, richly illustrated reports. You can easily log individual users or specific projects and compile detailed accounts of time spent within each program. The program helps track work time, pauses, projects, costs, software, and internet use and gives parents control over their children's use of the personal computer.

The software has some privacy safeguards and does not monitor all user data such as passwords and personal documents. Unlike spyware, it does not record specific keystrokes or run screen captures as a background process.

Software functions:

- User supervision: Get detailed accounts of working hours and breaks.
- Computer supervision: Monitor the family's PC or the company's network.
- Software metering: Determine how often software is used and by whom.

- Internet use: Control online time and web usage.
- Project overview: Summarize the amount of time applied to each project milestone.
- Compare users: See the activity of users on their computers.

Figure 6.1 Visual TimeAnalyzer in Action


Visual TimeAnalyzer - 220
File Edit View Help
Address Work - BDO
User
Monitor
Assistants
View Overview
Time period Today
from 14/08/2007 to 14/08/2007
Project »Work - BDO« : Overview of the last 1 days
Summary of times of day
PC on
Active
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 0 1 2o'clock
76% Work - BDO
7% Microsoft Office Outlook
7% Internet Explorer
2% Visual TimeAnalyzer
2% Microsoft Office Word
2% Security Task Manager
1% Openprocesserror
1% Windows Explorer
0% PaperMaster Pro 7.0
0% Microsoft Money
0% Microsoft Office Word
Summary of weekdays
17.0 h
Mon Tue Wed Thu Fri Sat Sun
Daily average
Work on PC: 17:00 h (119:00 h)
PC is online: 1:15 h (8:45 h)
PC is on: 24:00 h (160:00 h)
Number brackets is the weekly average
User
220
Programs
My projects
Education
Tax and Finances
Work - BDO
< Compare users >
Projects
View
Runtime
Last
Costs
Project time without PC use 12:50 h 1/07/2000 375 $
Microsoft Office Outlook 1:12 h 14/08/2007
Internet Explorer 1:11 h 14/08/2007
Visual TimeAnalyzer 0:21 h 14/08/2007
Microsoft Office Word 0:20 h 14/08/2007
Security Task Manager 0:18 h 14/08/2007
Openprocesserror 0:11 h 14/08/2007
Windows Explorer 0:09 h 14/08/2007
PaperMaster Pro 7.0 0:05 h 14/08/2007
Microsoft Money 0:05 h 14/08/2007
Microsoft Office Word 0:04 h 14/08/2007
PaperMaster Viewer 0:03 h 14/08/2007
StringFileInfo\040904B0\FileDescrip... 0:02 h 14/08/2007
Windows Calculator application file 0:01 h 14/08/2007
media service 0:01 h 14/08/2007

X-Ways Forensics

X-Ways Forensics also is covered in further detail in Chapter 6. It is based on WinHex and is designed to be an advanced work environment for the digital forensic analyst. The product provides a number of features such as:

- Forensic cloning and imaging of sound disks
- Examination of the complete directory structure inside raw image files, even if the directory spans segments
- Native support for FAT, NTFS, Ext2, Ext3, Ext4, CDFS, and UDF file systems
- Built-in interpretation of RAID 0 and RAID 5 systems and dynamic disks
- Viewing and dumping of physical RAM and the virtual memory space of running processes
- Several data recovery techniques and file carving
- File header signature database based on flexible UNIX grep notation
- Hard disk cleansing to produce forensically sterile media
- Gathering of slack space, free space, interpartition space, and generic text from drives and images
- Creation of file and directory catalogs for all computer media
- Easy detection of and access to NTFS alternate data streams (ADS), even where other programs fail
- Calculation for mass hash file encryption (CRC32, MD5, SHA-1, and SHA-256)
- Manual decompression of NTFS volumes
- Support for the HFS, HFS+, ReiserFS, Reiser4, UFS, and UFS2 file systems
- Support for the following partition types: MBR, Windows dynamic volumes, GUID (GPT), Apple, unpartitioned (floppy/superfloppy)
- Gallery view for pictures
- Calendar view
- File preview — a seamlessly integrated viewer component for more than 400 file types
- Examination of e-mail extracted from Outk (PST)★, Outlook Express (DBX), Mozilla (including Netscape and Thunderbird), generic mailbox (mbox, Berkeley, BSD, UNIX), Eudora, PocoMail, Barca, Opera, Forte Agent, The Bat!, Pegasus, PMMail, FoxMail, and local copies of maildir folders
- Automated check of the file signature
- Tagging of files and add them to customized report tables of notable items
- Automatic identification of encrypted Microsoft Office and PDF documents

- Automatic identification of pictures embedded in documents (for example Microsoft Word, PDF, and PowerPoint)
- Includes an internal viewer for Windows Registry files (all Windows versions) that generates an automated Registry report — works with all Windows versions
- Includes an internal viewer for Windows event log files
- Listing of the contents of archives within the directory browser, even in a recursive view
- Logical search of all or selected filesand directories, — following fragmented cluster chains, in compressed files, and optionally decoding text in PDF, WPD, and other file formats.
- Searches of Unicode and other types of code pages
- Skin color detection (that is, a gallery view sorted by skin color percentage that greatly accelerates searches for traces of child pornography and obscene images)
- Detection of host-protected areas (HPA) —ATA-protected areas

X-Ways Forensics can also write-protect data to ensure authenticity and integrity. It has a case management function integrated with automated activity logging (audit logs) and automated reporting. The generated reports can be imported and further processed by any application that processes HTML — for example, Microsoft Word. It also associates comments about files for inclusion in the generated reports or for filtering.

Evidor

Evidor searches for all occurrences of keywords within text files in digital media such as hard disks, and retrieves the context. It examines all files inthe entire allocated space, — including Windows swap/paging and hibernate files — along with unallocated and slack space. It will find data files that have been deleted if they still physically exist on the hard disk. Evidor is a small subset of the search functionality included in in X-Ways Forensics. Evidor cannot access remote networked hard disks.

Slack Space & Data Recovery Tools

These tools aid in the recovery of deleted files and data, including file fragments in slack space on file systems supported by Windows.

Ontrack

Ontrack Data Recovery is a simple-to-use tool used to recover lost, remote, and deleted data. It is provided with a file repair capability for files in Microsoft Word and Zip format. It also recovers deleted files, folders, and entire partitions. It uses an emergency boot diskette

to retrieve data from systems that cannot boot Windows. The user can start, stop, or resume the recovery process and select an FTP location as the destination for the recovered data. The user can configure the file filter for a quick or full scan.

The analyst can use this application perform an enhanced search using the options such as find, find next, and find previous. Ontrack can filter and sort data according to the file date, time, name, status, and size. It can scan the media for lost and remote data and lets the user specify the locations to copy the recovered files and folders to.

DriveSpy

DriveSpy is a modified MS-DOS shell designed to imitate and extend the potential of MS-DOS for forensic purposes. It uses a set of standard MS-DOS commands followed by commands specific to DriveSpy to process the computer during investigations.

DriveSpy can wipe an entire drive or partition, unallocated space, or slack space. It can also create an MD5 hash of an entire drive, partition, or selected files. It saves and restores compressed images of a partition for forensic use.

It can process the following:

- Hard drives with capacity of more than 8.4 GB
- Floppy drives and other removable drives
- FAT 12, FAT 16, /FAT 16x, FAT 32, and FAT 32x partitions
- Hidden MS-DOS partitions

Filters to sort the recovered files according to date, name, time, status, and size are available. DriveSpy performs the following functions for recovering data:

- Retrieve lost, remote, or deleted data and entire partitions.
- Start, stop, or resume the recovery process.
- Select an FTP location to copy the recovered files and folders to..

Data Recovery Tools

The following tools may be used to recover information from many sources including PDAs, cameras, and disk drives.

Device Seizure

Device Seizure is covert surveillance software developed by Paraben Corporation. It can aid in forensically analyzing and recovering mobile phone and PDA data including deleted text messages, photos, and call logs. It supports Palm, Windows CE, Windows Mobile, BlackBerry,

and Psion devices as well as certain Nokia, Sony Ericsson, Motorola, iDen, Siemens, LG, Samsung, and Symbian-based phone models. It also supports GSM SIM card acquisition and deleted data recovery using SIMCon technology. Device Seizure is supported on Windows 98, Windows XP, and Windows Vista.

Forensic Sorter

Forensic Sorter organizes and speeds the examination of the contents of a hard drive. It sorts the contents of hard drives into categories such as video, audio, and spreadsheets so you can easily find what you're looking for. It will filter common Windows files, recover deleted files, or file fragments in slack, deleted, and unallocated space.

The application supports drive images in RAW, PFR (Forensic Replicator), Safeback, and EnCase image files and is compatible with Paraben's P2 Forensic Examination Technology. It sorts files by header rather than file extension for accuracy and sorts files located in unpartitioned space. Forensic Sorter identifies encrypted files for easy recovery using Paraben's Decryption Collection. All data output is non-proprietary so any tool can be used for analysis performed after sorting. It is designed for use with Paraben's Case Agent Companion. It supports:

- EFS file detection
- NTFS compressed file read operations
- NTFS streamoperations
- HTML reporting
- Ext 2 and Ext 3 partition operations
- Compressed archive (ZIP, RAR, etc.) sorting
- Drives with bad sectors sorting
- Unicode support
- Removable disk support
- OLE storage support
- Palm OS file detection
- Chat log file detection
- E-mail Examiner and Network E-mail Examiner-supported file detection
- XML output validation

Directory Snoop

This program is a cluster-level search tool that allows Windows users to analyze and view files in FAT- and NTFS-formatted disk drives to see what data can be recovered. It can recover deleted files or permanently erase sensitive files. Supported media include hard drives, floppy disks, Zip disks, magneto-optical drives, and flashcard devices. Its features include:

- FAT and NTFS modules
- Recovering deleted files, including those emptied from the recycle bin
- Destroying sensitive files with the secure wiping functionality, which provides up to 35 wiping passes
- Secure wiping slack and free drive space
- Purging sensitive file names left behind after normal erasing
- Copying open and locked files with the cluster copy function
- Searching, filtering, and sorting files globally by name and other parameters
- Viewing, searching, printing, and copying raw cluster data
- Dynamically linking clusters with cluster chains and the files that use them
- Examining the FAT and Master File Tables
- Viewing files through external applications
- Verifying the effectiveness of other file wiping programs

Permanent Deletion of Files

Drive wiping is a crucial component of all digital forensic examinations. Any drive that is not thoroughly wiped has to be considered suspect. The following tools aid in this goal.

PDWipe

PDWipe is capable of wiping large hard drives (in excess of 8.4 GB) in just under 11 minutes. It will perform a declassification drive wipe in accordance with the NAVSO P5239-10 security standard. This wiping algorithm exceeds that specified by the DOD 5220.22-M specification for both "clearing" and "purging" of sensitive information on hard drives.

PDWipe provides the option of specifying a character code other than 0x00 when wiping a drive. It also offers the ability to wipe the drive using a random pattern. PDWipe can also record Logical Sector Addresses, and cylinder-head-sector (CHS) addresses for Int13 and Int13x geometries at the beginning of the sectors being wiped. This is useful when diagnosing architectural discrepancies when moving a drive between systems or validating imaging utilities.

PDWipe can process all drives in a system such that *all* drives can be wiped with a single program operation. If desired, PDWipe will generate a report of the wiping activity performed on a system. PDWipe can also verify that the contents of a specified number of randomly chosen sectors have been wiped. If wipe verification is requested, PDWipe will also automatically verify the first and last sector on the drive.

PDWipe will support any drive which accessible to your system through the Int13 or the Microsoft/IBM Int13 extensions. System BIOSs typically provide this capability for all attached IDE or EIDE devices. In addition, most SCSI adapters offer the ability to support attached devices through Int13 as well.

Darik's Boot and Nuke (DBAN)

Darik's Boot and Nuke (DBAN) is available from http://dban.sourceforge.net/ and is free. It is a self-contained boot floppy that securely wipes the hard disks of most computers. DBAN will automatically and completely delete the contents of any hard disk that it can detect, which makes it an appropriate utility for bulk or emergency data destruction.

DBAN can ensure due diligence in computer recycling, a way of preventing identity theft if you want to sell a computer, and a good way to totally clean a Microsoft Windows installation of viruses and spyware. DBAN prevents or thoroughly hinders all known techniques of hard disk forensic analysis.

File Integrity Checker

Failing to maintain the integrity of a file or drive image could be the end of a forensic examination. These tools help you to prove that the file you copied into evidence has not been altered subsequently. They make possible a quick and reliable diagnosis of a system image for the purpose of determining if any changes have occurred.

FileMon

Filemon and Regmon have been replaced by the Process Monitor on Windows 2000 SP4, Windows XP SP2, Windows Server 2003 SP1, and Windows Vista. Filemon and Regmon is included in these Windows versions for support of earlier Windows operating systems such as Windows 95.

FileMon monitors and displays file system activity on a system in real time. Its advanced capabilities make it a powerful tool for exploring the way Windows works, seeing how applications use the files and DLLs, or tracking down problems in system or application file configurations. Filemon's timestamping feature can show you precisely when every open, read, write, or delete operation happens, and its status column tells you the outcome. It begins monitoring when you start it, and its output window can be saved to a file for offline viewing. It has full search capability, and if you find that you're getting information overload, simply set up one or more filters.

File Date Time Extractor (FDTE)

File Date Time Extractor (FDTE) from Digital Detectives searches through binary files to uncover hidden, embedded, 64-bit date and time data. Windows files, such as Word documents contain several timestamp formats. It is important to note that false positives are common due to data patterns in files that look like timestamps but are not.

Decode - Forensic Date/Time Decoder

This utility was designed to decode the various date/time values found embedded within binary and other file types. It supports the following date/time formats and will allow you to specify the offset from GMT.

- Windows 64-Bit (little endian) date and time
- Windows 64-Bit (big endian) date and time
- Windows cookie-format date and time
- Windows filetime-format date and time
- UNIX 32-Bit (little endian) date and time
- UNIX 32-Bit (big endian) date and time
- UNIX numeric date and time
- MAC absolute date and time
- MS-DOS 32-Bit date and time
- HFS 32-Bit (little endian) date and time
- HFS 32-Bit (big endian) date and time
- HFS+ 32-Bit (little endian) date and time
- HFS+ 32-Bit (big endian) date and time

Date and time values are stored within Windows in various formats. For example, Internet History – (index.dat), recycle bin (INFO) files, Windows link files, and Microsoft Office documents contain a 64-bit date/time structure. Decode can take a decimal hexadecimal value and convert it into a date and time value in a variety of formats.

Disk Imaging Tools

These tools will create a bit-image copy of a drive or other media.

Snapback DatArrest

SnapBack DatArrest can obtain the mirror images of different operating systems. It provides successful backup of data present in the hard disk of a system. It is specially designed to create backup of mission-critical data backups.

The captured image contains all system and networking software with associated drivers, software applications, and configurations, and all the files including deleted files, slack space, and data files, as well as the CMOS settings for the system.

The data gathering takes place in several modes such as copying server hard drives to tape, PC hard drives to tape, server or PC hard drives to removable media, hard drives to hard drives, or tape to tape. It obtains a bit level backup of data present in a hard disk. The DatArrest Suite provides the ability to copy:

- Server hard drive to tape
- PC hard drive to tape
- Server or PC hard drive to removable media
- Hard drive to hard drive
- Tape to tape

DatArrest backs up any hard drive (SCSI, IDE, etc.) to any SCSI tape drive, removable drive, and even a SCSI or IDE hard drive.

Partition Managers: Partimage

Partimage is a Linux utility which saves partitions with a supported file system to an image file. Most Linux and Windows file systems are supported. The image file can be compressed with the gzip or bzip2 programs to save disk space, and they can be split into multiple files to be copied on CDs or DVDs, Partitions can also be saved across the network using the Partimage network support, or using Samba or NFS. If you don't want to install Partimage, you can download and run it from a CD.

Linux/UNIX Tools: Ltools and Mtools

LTools

The Ltools application is equipped with many command-line applications, which are executed from a MS-DOSwindow in Windows 9x, Windows ME, Windows NT, Windows 2000, or Windows XP. These applications provide the same functionality as the Linux "ls", "cp", "rm", "chmod", "chown", and "ln" commands. Therefore, the MS-DOS and Windows operating systems allows users to list Linux files and directories to copy files from Linux to Windows. Supported command sets include:

1. Delete or rename files in Linux (ldel and lren)
2. Creation of symbolic links (lln)
3. Create new Linux directories (lmkdir)
4. Modify a Linux file's access rights (lchange)
5. Change the Linux default directory (lcd)
6. Set the Linux default drive (ldrive)

Many of the tools present in UNIX operating systems have functions that are built into a single executable file. These functions are called a *bundle* of command-line parameters.

MTools

The MTools application is a group of tools that can allow the user to manipulate the UNIX systems. The MTools application can perform read, write, and other operations on MS-DOS files. Each application copies the commands that are present in MS-DOS. The MTools application also allows the user to unload and mount floppy disks.

The Coroner's Toolkit (TCT) and Tctutils

Dan Farmer and Wietse Venema developed The Coroner's Toolkit (TCT). This is a group of applications that can be used to aid a forensics investigation on UNIX systems. This system can be executed in most UNIX/Linux operating systems.

The Coroner's Toolkit has a collection of applications and plug-ins that give the program additional functionality. The Coroner's Toolkit was first put into use for the purpose of a postmortem analysis of a UNIX operating system after a break-in. The application was first given to a computer forensic class for real life applications in the field in 1999. This tool is mainly designed to help in a forensic investigation and is one of the most frequently used by forensic investigators.

Password Recovery Tools

A password cracker hashes all the words in a dictionary file and compares every result with the password hash. If a match is found, the password is the dictionary word. The following are tools that may be used to file poorly configured passwords.

@Stake

The LOphtCrack application has been upgraded to version 5, which is the latest version. The LOphtCrack application is an award-wining password auditing and recovery application used by many corporate companies worldwide. The LOphtCrack application uses many methods to reduce the security risk to network administrators by:

1. Identifying and resolving security vulnerabilities.
2. Recovering Windows and UNIX account password.
3. Rapidly processing accounts by using pre-computer password tables.

LOphtCrack equipped with many additional features which include automated password scanning that can be scheduled and Windows and UNIX operating system support. The application is also equipped with a remote scanning engine for multiple domains. There are many dictionaries for constructing new and secure passwords.

Decryption Collection Enterprise

Paraben's Decryption Collection Enterprise is an advanced password recovery suite with support for Windows Vista and Server 2003, EFS, SQL, and Lotus Notes along with support for everything included in the Standard Edition of the suite. One of its most advanced features is its ability to distribute its processing on a network of up to 16 computers. This allows brute force or dictionary attacks using several computers. Its features include:

- MD5 hash verification
- Easy-to-use interface - just drag and drop file to be recovered
- Listing of the most recent files recovered
- HTML reporting of recovery results
- Password cache for quick recovery of repeat passwords
- English password recovery accuracy of 90 percent and higher

AIM Password Decoder

This utility was designed to decrypt the login password for AOL Instant Messenger version 4. Please note that version 5 of AIM uses a different encryption method so this software will not work on that version.

MS Access Database Password Decoder

This utility was designed to decrypt the master password stored in a Microsoft Access database. There are two utilities in the ZIP file for decoding Access 95, Access 97, Access 2000, and Access XP.

FavURLView - Favorite Viewer

This utility will decode Internet Shortcut (URL) files to allow you to compare a shortcut description with the associated link. It will also decode the modified time and date. The software can be run as an external viewer within Encase, iLook, or any forensic application that supports external viewers. It has also been designed to accept data from Encase through the command line.

The software will also accept drag and drop multiple files - doing it this way allows you to copy out all of the URL files from your case and examine them all at once, thereby identifying any links that have been wrongly described in an attempt to hide the true link.

NetAnalysis

This product allows for the analysis of a Web browser's history data. It is commonly used by law enforcement in child pornography cases. The forensic examination and analysis of user activity on the Internet can be the pivotal evidence in any case. With the increase in the use of computers by pedophiles and other Internet criminals, it is vital for a forensic investigator to be able to extract this data and analyse it quickly and present the evidence in an understandable format.

NetAnalysis includes the following functionality:

- Viewing Web history, manually or automatically
- Viewing browser cache data
- extraction of browser data from unallocated space

The history extractor will also extract history records in a straightforward manner from a write-protected physical or logical device.

Multipurpose Tools

Maresware

This multipurpose tool is covered in Chapter 6 in more detail. It includes the following capabilities:

- Computer forensics on Windows and Linux
- Incident response
- Computer investigations
- Data analysis and auditing
- Drive wiping for reuse, retirement, or compliance with HIPAA and other information privacy regulations
- Data, software, and hardware validation tools
- Ownership identification of hard drives for inventory or for theft or loss recovery
- Systems maintenance and IT management
- Bates numbering of filenames for e-discovery

LC Technologies Software

PHOTORECOVERY by LC Technologies Software was developed for recovering images, movies, and media files from all digital media. It can be used with memory sticks, Smart Media, Connect Flash I and II, Micro Drives, SD and XD Cards, multimedia chips, floppy disks, and almost all other forms of digital film. Put the digital media in the external reading devices, run PHOTORECOVERY, and save the pictures in specified locations.

WinHex Specialist Edition

WinHex began as a disk-editing program and has developed into a forensic tool that is useful to digital evidence examiners of all skill levels. WinHex Specialist Edition can be used to validate the results of other tools and to perform specialized tasks such as file comparison and text extraction. This program can also teach novice analysts about disk layout, file system structures, data recovery, and other fundamental concepts in digital forensics.

WinHex is a primarily a hexadecimal editor. It inspects and edits all types of files and can also recover deleted files and lost data from hard drives. It can do this on hard drives with corrupt file systems or from digital camera cards. It includes a RAM editor that provides access to the virtual memory of other processes. It is useful in concatenating and splitting files and unifying and dividing odd and even bytes and words. It also has the ability to analyze examine, and compare files. It can create a drive image and also make backups of drives. It can import all clipboard formats including ASCII hexadecimal values.

Prodiscover DFT

ProDiscover DFT reads the disk at the sector level, circumventing the standard file system while it recovers files, looks in slack data,the HPA section of a disk, and examines alternate data streams in Windows systems. This unique approach also allows you to examine files without altering valuable metadata such as the last-time-accessed date.
The main features include:

- Conforms to NIST's disk imaging tool specification 3.1.6 to ensure high quality
- Creates a bitstream copy of the evidence disk
- Supports FAT12, FAT16, FAT32, and NTFS file systems
- Analyzes data in a protected area (ATA) by identifying the partitions in the protected area
- Calculates the file checksum using either SHA1 or MD5
- Analyzes data in alternate data streams (ADS) of NTFS partitions
- Capable of analyzing dynamic disks

Toolkits

NTI Tools

NTI Tools offers a compilation of software forensic tools. It is available from www.forensics-intl.com/tools.html. The products in this tool set include:

Stealth Suite

Users without a forensic background can use the Stealth Suite to assess activity on a computer hard disk. These tools can help identify whether a targeted computer system was used to access inappropriate information.

Computer Incident Response Suite

This suite of tools are often used in corporate and government investigations and security risk reviews. They are optimized for MS-DOS, which is the lowest cost forensic platform for MS-DOS and Windows processing. Many of the tools also have Windows version. This should be one of your first forensic tool sets. It also makes an excellent set of tools to cross-validate your findings before you go before the court or the board.

Data Elimination Suite

This suite of tools performs computer data elimination and validation. It allows you to remove information from a drive and cross-validate that the information has been removed. It eliminates classified data 'leakage' and verifies that the data was properly eliminated.

TextSearch Suite

NT is used to process Windows NT/2000/XP-based computer systems from a MS-DOS command line. TextSearch Suite provides the same interface and features as its predecessor, TextSearch Plus, and also identifies many compressed and graphics files using the file header signature. giving the investigator a listing of files that could store information in a compressed or graphic format.

The HexSearch utility is also included in this suite. This tool provides a similar interface as TextSearch Plus while allowing the user to search for hexadecimal strings, such as file headers, non-printing characters, and more.

NTI Secure ToolKit

This software is used to secure sensitive files stored on portable and desktop computers. Because it uses NIST- tested and approved AES 256-bit encryption, it qualifies for government use with classified *'Secret'* level data. This software exceeds commercial security requirements and it is much easier to use than Pretty Good Privacy (PGP). It includes a management tool so that corporate information is not lost to the corporation. An export license may be required for locations outside the United States.

SafeBack 3.0

This makes evidence grade bitstream backups of hard drives. NTI incorporated the NIST approved 256-bit SHA algorithm to provide greater protection against manipulation.

R-Studio

R-Studio is a family of powerful and cost-effective data recovery software. It recovers files from FAT 12, FAT 16, FAT 32, NTFS, NTFS 5 (created or updated by Windows 2000, Windows XP, Windows 2003, or Windows Vista), HFS/HFS+ (Macintosh), little- and big- endian variants of UFS1/UFS2 (FreeBSD, OpenBSD, NetBSD, and Solaris), Ext2FS (Linux), and Ext3FS (Linux) partitions. It functions on local and network disks, even if such partitions are formatted, damaged, or deleted. The suit includes the following tools:

- R-Studio
- R-Undelete
- Drive Image
- R-Firewall
- R-Guard
- R-Mail
- R-Word
- R-Excel
- R-Wipe&Clean
- R-Linux

Datalifter

DataLifter comes in two versions, DataLifter version 2.0 and DataLifter.Net Bonus Tools. Two additional packages, DataLifter - Digital Companion and DataLifter - File Extractor Pro, are add-ons. DataLifter - File Extractor Pro includes the following features:

- Multiple-thread data carving operations
- Support for source files larger than 2 GB
- Support for compressed and uncompressed Encase evidence files
- Support for logical disk access for sector-by-sector searches or searches on unallocated clusters
- An automated File Signature Builder that scans target systems so each data recovery system can be customized for the target data

- File recovery is completed through multi header points and footer values
- Twenty-three pre-configured file signatures
- File signature selection is easily organized by custom categories
- Organized recovery output
- Pausing and restarting the data carving session
- Create MD5 hash of each extracted file allowing optional deletion of duplicate files, saving you time during review
- Keyword support as a condition for reducing false positives
- Portable — can be registered and run from a thumb drive

Forensic Toolkit (FTK)

The Forensic Toolkit (FTK) is a commercial forensic software package by AccessData. FTK can parse FAT12, FAT 16, FAT 32, NTFS, NTFS Compressed, Ext2, and Ext3 file systems. It can use image files created by EnCase, SMART, Snapback, some versions of Safeback, and dd. The program allows users to search with keywords or take advantage of drive indexing using the dtSearch algorithm.

Image Master Solo and Fastbloc

Image MASSter Solo

MASSter is used as a system integration and MIS tool for system integration and system rollouts. It is widely used for hard disk duplication. It is recommended for creating image drives. The following describes the capabilities of MASSter:

- It performs 100 percent copying of data to disk drives.
- The data transfer rate is 800 Mb/minute.
- Parallel transfer rate is 40 MB/minute.
- A sample audit trail report is available on-site with an option of a serial printer.
- It is functional for a variety of devices on the market.
- It also performs the job of wiping out unnecessary data.
- It ensures data integrity.
- MASSter is used to capture the images of affected system screens that can reveal evidence about the crime.
- It performs disk-cloning activities.

- It can perform high-speed data transfer from the affected computer to any other device.
- "Wipeout" functionality erases and sanitizes the drive you are copying.
- It comes with some external devices such as CD Writer, Serial Printer, and Battery Pack for Printer.

FastBloc

FastBloc LE (formerly FastBlocIDE) is a data acquisition and write protection device. It installs on any PC with a spare 5.25-inch drive bay.

- 40-pin IDE interface connection
- Host controller connection
- 40-pin & 44-pin IDE connection
- Suspect drive connection
- ATA-6 support
- An ATA-6-compliant host controller

FastBloc LE is a reliable solution for the forensic lab environment. FastBloc LE can be placed in a 5¼-inch external outlay of the forensic computer. A connection is established between FastBloc LE and other devices using a direct IDE-to-IDE device so the data can be acquired remarkably fast and does not require any other complicated or expensive hardware. This tool can be used to acquire drives without physically opening the forensic computer. When used with an ATA 6 controller, it can transfer data up to a rate of 133 MB/s and can also support drives bigger than 128 GB.

The main features of FastBloc are:

- It makes use of IDE controllers and doesn't require SCSI controller card drivers.
- It has similar height and width as the internal CD-ROM drive.
- Small size (5.75" W × 3" D × 1.625" H) allows for complete portability and ease of use in the field.
- Has ports for desktop IDE as well as laptop IDE media.

Encase

Guidance Software states that, "*EnCase provides a much simpler way to conduct a search of a computer system, document the findings, and make evidentiary and discovery copies*". EnCase is a forensic data and analysis program for Windows 95 and NT that is designed to provide case management and tools for digital investigations. EnCase enables the user to understand and define the information in the system. The information collected in EnCase can be exported to several file formats.

E-mail Recovery Tools

Paraben's E-mail Examiner

This product provides forensic analysis, advanced searching, and converting and Exporting of e-mail. E-mail Examiner can examine over 16 e-mail formats, including AOL 9.0, PST files, and more than14 others. It includes the following features:

- Support for files larger than 2 GB
- Handling of bad date fields in message headers
- Handling of plain UTF-8 messages and header fields
- Handling of corrupted DBX files
- Parsing of HTML mail
- Support for maildir files
- Support for Opera 9 .mbs files
- Support for Agent 3.x and 4.x folder name conversion
- Aid4Mail console for importing, exporting, and converting
- AOL and PST Converters
- Export any e-mail format to PST
- Export to MHTML with XLS or database index
- Recovers deleted/deleted e-mail
- America Online (AOL) 9.0
- Outlook Exchange (PST)
- USENET Groups
- Eudora
- Netscape Messenger
- Pegasus Mail
- Outlook Express
- The Bat!
- Forte Agent
- PocoMail
- Barca

- Calypso
- FoxMail
- Juno 3.x
- EML message files
- Mozilla Mail
- MSN Mail
- Generic mailboxes (mbox, , UNIX mail format)

Network E-mail Examiner

Network E-mail Examiner allows the examination of Microsoft Exchange (EDB), Lotus Notes (NSF), and GroupWise e-mail stores. Network E-mail Examiner is designed to work hand-in-hand with E-mail Examiner and all output is compatible and can easily be loaded for more complex tasks.

The product features include:

- Support for GroupWise information stores up to 6.5.1
- Support for Microsoft Exchange 5.0, 5.5, 2000, and 2003 SP1 information stores
- Support for Lotus Notes information stores 4.0, 5.0, 6.0
- Advanced search options
- View one or all individual e-mail accounts in the information store
- View all metadata within individual messages
- Complete bookmarking functionality
- Export data into Paraben's E-mail Examiner
- Summary HTML reporting
- Export to PST files
- Output to MSG & EML format
- Supports deleted/deleted recovery with Exchange

Oxygen Phone Manager

Oxygen Phone Manager II for Nokia phones provides a simple and convenient way to control mobile phones from a PC. Oxygen Phone Manager II offers management for the phonebook, call register, calendar, to-do lists, SMS and MMS messages, logos, tones, GPRS and WAP settings, profiles, dictaphone, FM stations, Java games, and applications.

SIM Card Seizure

Paraben acquired SIMCon and integrated it into SIM Card Seizure and Device Seizure. SIM Card Seizure can be used to recover deleted Short Message Service (SMS) messages and perform comprehensive analysis of SIM card data. It is a combination of the SIM card acquisition and analysis components in Paraben's Device Seizure utility and a specialized SIM card forensic acquisition and analysis tool. SIM Card Seizure includes the software and a forensic SIM card reader. Its features are included in Device Seizure and the Device Seizure Toolbox.

SIM Card Seizure has Unicode support for languages such as Arabic, Chinese, and Russian, and includes the following features:

- Forensic SIM Card reader
- Calculates MD5 and SHA1 Hash Values
- Search function
- Recovers deleted SMS data
- Bookmarking options
- Report Creation wizard
- Save Workspaces for Further Review functionality
- Timestamps Calculate GMT Offset functionality
- Access to Paraben's forum
- Access to Paraben's 24-h our support
- Data acquired from SIM cards
- Phase ID
- SIM Service Table (SST)
- ICCID serial number
- Preferred Languages (LP) variable
- Service Provider Name (SPN)
- MSISDN subscriber phone number
- AND short dial number
- FDN fixed numbers
- Last Dialed Numbers (LND)
- EXT1 Dialing Extension
- EXT2 Dialing Extension

- GID1 Groups
- GID2 Groups
- SMS text messages
- SMSP Text Message parameters
- SMSS text message status
- CBMI preferred network messages
- PUCT charges per unit
- ACM charge counter
- ACMmax charge limit
- HPLMNSP HPLMN search period
- PLMNsel PLMN selector
- FPLMN forbidden PLMNs
- CCP capability configuration parameter
- ACC access control class
- IMSI
- LOCI location information
- BCCH broadcast control channels
- Kc ciphering key

Autoruns

This product is available from Microsoft at: www.microsoft.com/technet/sysinternals/Utilities/AutoRuns.mspx.

This utility has the most comprehensive display of autostart locations of any startup monitor and displays all programs are configured to run during system boot or login. It displays the entries in the order Windows processes them. These programs include those in your startup folder, Run, RunOnce, and other Registry keys. You can configure Autoruns to show other locations, including Explorer shell extensions, toolbars, browser helper objects, Winlogon notifications, autostart services, and much more. Autoruns goes beyond the MSConfig utility bundled with Windows Me and XP.

Autoruns' Hide Signed Microsoft Entries option helps you to zoom in on third-party autostart images that have been added to your system and it has support for looking at the autostarti images configured for other accounts configured on a system. It also includes a command-line equivalent that can output in CSV format — Autorunsc.

HashDig

HashDig technology is a collection of utilities designed to help analysts automate the process of resolving MD5 and SHA1 hashes. The product site states, "*In the early stages of an investigation, it is not typically possible or practical to examine all subject files. Therefore, practitioners need reliable methods that can quickly reduce the number of files requiring examination. One such method is to group files into two general categories: known and unknown. This method can be implemented quite effectively by manipulating hashes and comparing them to one or more reference databases. Even that, however, can take a significant amount of effort. HashDig technology attempts to reduce this burden through automation and the use of lightweight, open, and verifiable techniques*".

Patchit

Created for software developers, this is a two -art program used to make patches for updating software. Both the manager and the patch program are simple-to-use GUI programs. Command-line switches are also available.

PowerGREP

PowerGREP is a powerful Windows grep tool. It will search through large numbers of files on a system or network, including text and binary files, compressed archives, MS Word documents, Excel spreadsheets, and PDF files. This is done through the use of powerful text patterns (regular expressions) by specifying the form of what you want instead of literal text. It will search and replace with one or many regular expressions to comprehensively maintain web sites, source code, reports, and other sources. PowerGREP will also extract statistics and knowledge from logfiles and large data sets.

Reverse Engineering Compiler

REC is a portable reverse engineering compiler, or decompiler. It reads an executable file and attempts to produce a C-style representation of the code and data used to build the executable file. It is portable because it has been designed to read files produced for many different targets, and it has been compiled on several host systems.

RecStudio offers a modern user interface to REC's interactive mode. A command-line version is still available for Linux and Solaris hosts.

These are some of REC's features:

- Multitarget operation: REC can decompile 386, 68k, PowerPC, and MIPS R3000 programs.
- Multiformat operation: REC recognizes the following file formats:

- ELF (System V release 4, for example Linux and Sun Solaris)
- COFF (System V release 3.x, for example SCO)
- PE (Win32.exe and .dll for Windows 95 and NT)
- AOUT (BSD derivatives, for example SunOS 4.x)
- Sony PlayStation PS-X (MIPS target only)
- Raw binary data (thorough CMD files)

- Multihost operation: REC is available for Linux 3.0 (i386), Windows 95, and Sun SunOS 4.1.4.
- Supports high-level symbolic information in COFF, ELF+STAB, and AOUT+STAB.
- Scalable user interaction: From 100 percent batch mode to a full-screen, browser-like interactive mode.
- HTTP server mode allows the use of an HTML browser as the user interface.

Table 6.1 contains a list of forensic software tools.

Table 6.1 Forensic Software

AOL PW Extractor	Reveals all AOL passwords on your computer.
Accent Access Password Recovery	Recovers forgotten or lost passwords for Microsoft Access documents.
Accent Excel Password Recovery	Recovers forgotten or lost passwords to open for Microsoft Excel documents.
Accent Money Password Recovery	Recovers forgotten or lost passwords for Microsoft Money documents.
Accent Office Password Recovery	Recovers forgotten or lost passwords for Microsoft Office documents.
	If you have lost or forgotten a password for opening a Microsoft Office document (Access, Excel, and Word), or a password for saving changes to a Microsoft Word document, you can use this software. It will help you search for a password more effectively. The software offers you to choose one of the three ways: a dictionary-based attack, a brute force attack or a brute force attack with a mask.

Continued

Table 6.1 Continued

ACDSee	Fast photo viewer and manager. Easily find, view, manage, print, edit, and share images.
Active Partition Recovery	Recovers deleted partitions (FAT and NTFS). Restores deleted FAT and NTFS Logical Drives. Creates –an image of a drive for backup purposes. Scans hard drives and detects deleted FAT and NTFS partitions or logical drives. Preview files and folders on a deleted partition or drive to recover data.
	A very small, easy to use MS-DOS program (only 150k in size)
Active Ports	Shows all open TCP/IP and UDP ports on Windows NT, Windows 2000, and Windows /XP computers and maps them to the parent application.
Active UNDELETE	Restores deleted files and directories.
Active Uneraser for MS-DOS	Data recovery (undelete) software for MS-DOS. Recovers deleted files on FAT16, FAT32, and NTFS partitions.
Adobe Reader	PDF reader.
Advanced Attachments Processor	Extracts attachments from e-mail client message databases, and makes an archive of the extracted files.
Advanced Email Extractor	Extracts e-mail addresses from web-pages on the Internet (using HTTP and HTTPS protocols) and from HTML and text files on local disks.
Advanced Mailbox Processor	Extracts owner's names and e-mail addresses from local files to creates a list.
Advanced Forensics Format (AFF)	An extensible open format for the storage of disk images and related forensic information.
Afind	Afind lists files by their last access time without altering the data the way that Windows Explorer does . Afind searches for access times between specified time frames, and coordinating the results with logon information provided from ntlast. This enables you to determine user activity even if file logging has not been enabled.
Automated Image and Restore (AIR)	A GUI fron end to dd/dcfldd creates forensic bit images.
Allin1 for Sleuthkit	Runs time consuming tasks in Sleuthkit/autopsy in batch mode.
APDFPRP	PDF file cracker.
ASP	Zip file cracker.

Table 6.1 Continued

Autopsy Forensic Browser	HTML front end for TCT and TCTUtils. It has a file manager style interface, can display the contents of a file as raw data or in ASCII, and generates reports.
AutoStart Viewer	When you start Windows, dozens of programs are already running – many of them invisible and running in the background. This software identifies what is running, why it is running and determine if any are Trojans. AutoStart Viewer allows you to see every AutoStart on your system on one screen. In addition, it gives you complete control over the AutoStart references, and allows you to modify or delete them at will.
AVIPreview	A program which is capable of playing partially downloaded AVI movies.
Back2Life	Simple undelete utility for Windows
BadCopyPro	Data recovery software for floppy disk, CD, DVD, memory card, Zip disk, flash drive, and other storage media.
BCWipe	Disk wiper (DoD – 7-pass wipe tool)
Belkasoft Forensic IM Extractor	Supports various IMs: ICQ versions 99a up to ICQ5, MSN Messenger, Yahoo! Messenger, &RQ, Miranda. Supports deep ICQ analysis using different methods (with and without usage of index file) that allows user to extract even deleted and overwritten messages.
BinText	Finds ASCII, Unicode, and resource strings in a file.
BitForm Discover	Analyzes and reports information about metadata and hidden information within document collections.
Black Bag Macintosh Forensic Software	A suite of forensic solutions and a Macintosh boot CD that boots any system capable of running Mac OS X
bmap	Stores data in slack space on ext2 filesystems, among other features. Forensic examiners can use bmap to detect used slack space and to recover data.
Browser History Viewer	Examines the contents of web browser history files and exports the data.
BXDR	This utility will display the full sector count on a hard disk drive, including protected areas that utilities such as Safeback and Encase miss when imaging.
ByteBack	Data recovery and forensic software program that accesses media at the physical level.
Cache Reader	Reads the index.dat file in the Temporary Internet Files (TIF) folder of Internet Explorer versions 5 and 6.

Continued

Table 6.1 Continued

Cache View	Opens cached files for viewing and copy or move them out of the cache. This will even reconstruct the names and directory paths of Web sites so you can view the HTML files. This supports all versions of Netscape, Mozilla (Gecko), and Internet Explorer.
CacheInf	Displays a browser's cache, and provides the option to delete it. Additionally, searches the cache for file names and URLs and saves the search results into a comma-separated value file.
CacheView	Cache View is a viewer for the Netscape Navigator, Mozilla, and Internet Explorer caches. It displays cached files , gives you the option to copy them or move them out of the cache. It can reconstruct the names and directory paths of the files. Cache View extracts the following information about cached files: URL, size (in bytes), MIME type, last modified date, the date the file was downloaded, and the expiration date.
CacheX for IE	CacheX displays a browser cache with a Windows-Explorer-like user interface. To view a file offline with your browser, double-click it.
Captain Nemo	Connects a drive containing a UNIX/Linux , Windows NT, or Novell operating system directly to a computer running the Windows operating systemfor accessing, viewing, printing, and copying the files as if they were on another Windows drive on the computer. This product only supports ext2 file systems on Linux.
Captive	The first free NTFS read/write file system for GNU/Linux.
Catalogue	File Metadata Miner; a file cataloging utility that enables quick creation of HTML page listing files and associated metadata, managing and updating document properties associated with such files
CD Roller	Retrieves data from a hard disk created by "drag and drop" CD and DVD writing software, such as Roxio (Adaptec) and Ahead's Nero, CeQuadrat's PacketCD, Instant Write, and B's CliP.
CD/DVD Inspector	Analyzes and extracts data on CD-R, CD-RW and DVD media. Tailored for use by professionals in data recovery, forensics, and law enforcement.
CD Roller	Toolset for CD and DVD data recovery.

Table 6.1 Continued

Chip-it	Extracts phone numbers from a variety of mobile phones – freeware.
chrootkit	Checks for evidence of modification of the system binaries by a rootkit on a local machine.
CMOS Recovery Tools	Recovers CMOS passwords. Works with the following BIOSs: ACER/IBM, AMI, AMI WinBIOS 2.5, Award 4.5x/4.6x, Compaq (1992), Compaq (new version), IBM (PS/2, Activa, Thinkpad), Packard-Bell, Phoenix 1.00.09.AC0 (1994), Phoenix a486 1.03, Phoenix 1.04, Phoenix 1.10 A03, Phoenix 4.05 rev 1.02.943, Phoenix 4.06 rev 1.13.1107, Phoenix 4 release 6 (user), Gateway Solo, Phoenix 4.0 release 6 , Toshiba, Zenith AMI.
CookieView – Cookie Decoder	Decodes the internal cookie data such as the date and times and splits the data into separate cookie records. This software was originally written as an external viewer for Encase or iLook. Either drag and drop a cookie onto the main window or set it as an external viewer.
Coreography	Browses memory images. Coreography is an open-source utility.
Creed (Cisco Router Evidence Extraction Disk) - Raw Image (Freeware)	Creates andrestores disk images, even non-PC formatted disks. This software was created by New Technologies Inc. Once you have downloaded the zip file, unzip it into its own directory. This is a MS-DOS application, so no further installation is required.
Crucial ADS	Finds hidden data streams in NTFS file systems.
Digital Investigation Manager (D.I.M.)	Manages incident response and forensic acquisition procedures. D.I.M. allows operations to be organized by case. Each case may contain an unlimited number of hosts (for example, workstations, servers, laptops, and PDAs). Items of evidence (hard disks, CD/DVD-ROMs, memory card, log file, network dumps) are associated with each host .
DataGrab	Investigates Internet Relay Chat (IRC) communications.
Davory	Davory recovers files from logically damaged or formatted drives. Created by the makers of WinHex.
DbExtract	Extracts mail messages from Outlook Express version 5 DBX files. It requires the existence of the Visual Basic 6 runtime DLL — msvbvm60.dll.

Continued

Table 6.1 Continued

DBXanalyzer	Reads, analyses, and manages e-mail data files created by Microsoft Outlook Express versions 5 and 6.
DBXpress	DBXpress is a faster, more accurate, and more powerful version of DBXtract. It requires that .NET Framework version 1.1 be installed.
DBXtract	DBXtract extracts all e-mail and news messages from individual DBX files
DCFL-DD	An enhanced dd with MD5 hashing.
dd for Windows	Flexibly copies data within in a Win32 environment.
dd rescue	Copies data from one file or block device to another.
DecExt	Recovers base 64 pictures.
Declasfy	Wipes drives according to DOD specifications. Drive wiping with Declasfy can serve many purposes where information security is a concern. For example: preparing drives for internal reuse; securing private information prior to retirement or donation of a drive; securing private information for compliance with HIPAA and other regulatory requirements.
Decode	Decodes the date/time values embedded within binary and other file types.
DFSee	Display File Systems (DFSee) is a generic partition and file system utility. It supports the following partition tables: FDISK, LVM, (V)FAT, FAT-32, HPFS, NTFS, and JFS.
Digital AudioRescue Professional	This program recovers lost data from multimedia devices including digital audio recorders, MP3 and WMA players, PDAs, and mobile phones. This program supports recovery from hard drives, CompactFlash cards, IBM Microdrives, SmartMedia cards, MultiMedia cards, Secure Digital cards, and Memory Sticks.
Digital Image Recovery	Reconstructs deleted images, videos, and audio files from formatted media, even if the media was ejected during a write process.
Directory Snoop	Snoops FAT and NTFS formatted disk drives to see what data may be hiding in the cracks. This is a cluster-level search tool. Use Directory Snoop to recover deleted files or permanently erase sensitive files. Supported media includes local hard drives, floppy disks, Zip disks, MO disks, and flashcard devices.

Table 6.1 Continued

DIRV	A filter for the DIR /S command. Dirv is a program for those who still use DIR /S to obtain a recursive directory list of all the files on a system. The DIR/S program produces an output that is difficult to import into a database for additional processing. Dirv takes the output generated on file systems supported by Windows NT or Windows 9x and converts it to one-line records that contain the appropriate path and filename merged.
Disable	Disables a computer keyboard.. Best used on a boot disk for evidence protection, which isften called an evidence disk.
Disk Investigator	Disk Investigator helps you to discover all hidden data on a hard disk.
DiskCat	Catalogues all files on disks. DiskCat is short for "disk cataloguer". It creates a listing (catalogue) of all filesor directories on a hard or floppy disk. It is customizable and is especially useful for forensic purposes and file maintenance. Its output is a fixed-length record and database compatible for further analysis and sorting.
DiskPatch	Repairs Master Boot Records (MBRs), partitions, and boot sectors. It also clones and wipes disks and scans for errors.
disktype	Detectsthe format of a disk or disk image. It detects common file systems, partition tables, and boot codes.
DivX Player	The DivX codec plays any DivX video data,including DivX VOD movies.
DNA Manager	Distributed network password cracker from Access Data. Harnesses networked PCs to do this.
DriveLook	DriveLook scans a drive or partition for text strings and stores them in a table. You can browse this table and view the locations where the words had been found. The search function enables fast inquiries for combinations of words. The program also enables you to index all of the text on a hard drive, browse a list of all words stored on the drive, search for words or combinations of words, view the location of words in a disk editor, switch between several views(such as hexadecimal and text), use physical drives or logical drives as an input, use image files as an input, and access remote drives over a serial cable or TCP/IP connection. DriveLook is used for forensic drive investigation and data searches.
dtSearch	Provides over two dozen indexed and unindexed text search options for all popular file types.

Continued

Table 6.1 Continued

dumpautocomplete	Dumps Firefox AutoComplete files into XML files.
e2recover	Undeletes data on ext2 file systems.
e2Salvage	Recovers lost ext2 partitions.
Eindeutig	Parses Outlook Express DBX files.
E-Mail Detective	Extracts all email data (including graphics) from America Online's databases stored on a disk drive.
EMF Printer Spool File Viewer	Displays the contents of an EMF spool file.
Encase 4	Performs most of digital forensic analysis operations.
Ethereal	Analyzes network packets. A network packet analyzer will try to capture network packets and display that packet data in as much detail as possible. Data can be captured "off the wire" from a live network connection or read from a capture file. It runs on all major operating systems, including UNIX, Linux, and Windows. Live data can be read from the following network types: Ethernet, FDDI, PPP, Token-Ring, IEEE 802.11, Classical IP over ATM. Loopback interfaces are also supported. However, not all of these network types are supported on all operating systems.
Event Analysis and Reconstruction in Lisp (EARL)	Reconstructs finite state machine events. This is experimental software.
Event Log Explorer	Views, monitors, and analyzes events recorded in the security, system, application and other logs produced by the Windows operating system.
Eventlog	Takes the output of an Windows NT security eventlog and reformats it to single lines containing pipes. These can be easily imported into a database or spreadsheet.
Evidence Mover	Automates the transfer of evidentiary data from one location to another with 100 percent accuracy.
Evidor: The Evidence Collector	Retrieves the context of keyword found on computer media, Windows swap/paging files, hibernate files, and unallocated space and slack space. Created by the makers of WinHex.
EXIF Image Viewer	Reds EXIF information embedded in photos and thumbnails.
EXIF Reader	Analyzes and displays the shutter speed, flash condition, focal length, and other image information included in the EXIF image format. The EXIF image format is supported in many digital cameras.

Table 6.1 Continued

Exifer	Recovers and displays the metadata (EXIF/IPTC) included in pictures taken by digital cameras.
EXIFRead	Utility that extracts image information from EXIF/JPG files.
Explore2fs	The Win32 explorer for Linux ex2fs partitions; reads ex2fs filesystems under NT 4.0 (and Windows 95). It is a seperate program - not a file system driver.
Extract 2.10	Extracts files from a disk image created using WinImage, FDFormat, or compatible tools running under MS-MS-DOS, Windows 95, Windows 98, or the Windows NT console.
Fatback	Undeletes files from FAT file systems.
faust	Analyzes files found after an intrusion or the compromising of a honeypot. faust is a perl script.
File Date & Time Extractor (FDTE)	Searches binary files for hidden, embedded, 64-bit date and times.
File Disk	A virtual disk driver for Windows NT, Windows 2000, and Windows XP that uses one or more files to emulate physical disks. A console application is included for dynamically mounting and unmounting files.
File Investigator	Identifies files by content and extensions. 1,806 file formats are supported.
File Juicer	Extracts images from binary files on the Macintosh.
File Scavenger	Undeletes files on NTFS volumes. Runs on Windows NT, Windows 2000, and Windows XP. Reformatted volumes or broken hardware/software RAID volumes can also be recovered.
FileSystem Investigator	Views and extracts file system data. Platform independent.
Final E-mail	Scan Outlook Express, Eudora, and Netscape Mail e-mail database files and recovers messages. Also, locates lost emails that do not have data location information associated with them
FINAL Photo Retriever	Recovers image files, video files, and audio files from hard disks and removable media.
FinalData	Recovers data. Runs on Windows NT 4.0, Windows 2000, Windows XP.
FINALdBase 2.0 for Oracle	Recovers Oracle databases. Runs on UNIX systems.

Continued

Table 6.1 Continued

FinalEmail	Recovers the e-mail database file and locates lost e-mails that do not have data location information associated with them.
Forensic and Log Analysis GUI (FLAG)	Analyzes log files for use in forensic investigations.Often, when investigating a large case, a great deal of data needs to be analysed and correlated. FLAG uses a database as a back-end to assist in managing the large volumes of data.
Flash Retriever Forensic	Recovers pictures, movies, and sounds from a variety of media.
Floppy Image	Create image files of floppy disks and back (for backup, shipping or transfer). Save the image file compressed, uncompressed or as a self-extracting exe. Add descriptions to or convert your old image files.
Foremost	Recovers files based on their headers and footers. Runs on Linux systems. Foremost can process image files, such as those generated by dd, Safeback, Encase.
Forensic Acquisition Utilities	A collection of utilities and libraries used in forensic-related investigations.
Forensic Analysis Toolkit (FATKit)	Analyzes volatile system memory. A cross-platform, modular, and extensible digital investigation framework.
Forensic Internet Explorer – Beta	Parses the "index.dat" file, which identifies host sites that have been visited.
Forensic Script	Modified version of a script from 'Forensic Analysis of a Live Linux System, Part One' by Mariusz Burdach, as well as a server-side script for netcat, and description of the relationship between a specified set of ports and files.
Forensic Tools on the Mac (by M Dornseif)	Compiles various forensic tools on the Macintosh.
Foundstone's Forensic Toolkit	Examines NTFS for unauthorized activity.
FSDEXT2	Transparently mounts Linux ext2fs partitions on Windows 9x systems.
FS-TST	Tests disk imaging tools typically used in forensic investigations. This package includes programs that use the Int13 BIOS disk interface to initialize disk drives, detect changes in disk content, compare pairs of disks, and simulate bad sectors on a disk.

Table 6.1 Continued

FTimes	Creates system baselines and assists in collecting evidence. Its primary purpose is to gather or develops information about specified directories and files in a manner conducive to intrusion analysis.
AccessData Forensic Toolkit (FTK)A complete forensic toolkit.	
ftrace	Performs fast traceroutes Win32-based systems.
Galleta	Reconstructs a subject's Internet Explorer cookie files.
Gargoyle Forensic Pro	Determines whether malware is present on a system under investigation.
Gemulator Explorer	Reads Atari ST- and Apple Macintosh-formatted disks. Runs on Windows systems.
Symantec's Norton Ghost 9.0 Performs advanced backup and recovery operations.	
GNU utilities for Win32	Native Win32 ports of several GNU utilities.
gpart	Attempts to determine the primary partition table of a PC hard disk when the primary partition table in sector 0 is damaged, incorrect, or deleted. If identified, the table can be written to a file or device.
GRAB	Acquires image data. Runs on Linux systems. Created by the developers of Helix.
Handle	Displays information about open handles for any process in the system. You can use it to see the programs that have a file open, or to see the object types and names of all the handles allocated by a program.
Hard Disk Copy	Backs up and copies hard disk sectors.
Hard Drive Mechanic Deluxe	Repairs hard disk drives.
HD95Copy	Copies a hard drive to an image file on a network server, another hard disk, or any medium you can access using a logical drive letter. The hard disk is copied sector by sector, provided that filenames and hidden files and directories are backed up also. It also supports removable media.

Continued

Table 6.1 Continued

HD98Copy	Same functionality as HD95Copy. Documentation in German and English
hfsutils	Manipulates HFS volumes on UNIX and other systems. HFS is the "Hierarchical File System," the native volume format used on modern Macintosh computers.
History Inspector	Reads a Internet Explorer versions 5 and 6 history database (index.dat) file and presents it as a synoptical table. It also supports link browsing, adding Favorites, copying URLs, adding personal notes, searching and replacing text items, and printing.
History Inspector for Internet Explorer	Reads all of the information in the history database and presents a list in chronological or alphabetical order.
Host Protected Area (HPA)	Identifies the following information about a disk drive: the manufacturer, serial number, total number of sectors on the drive, and the number of sectors set aside in the (Host Protected Area (HPA) if one exists on the drive. HPA is very useful on a forensic boot disk because it can capture key information about any IDE drives in the system. This information can be sent to an output log file for future reference.HPA is a 16 bit program designed to work only on IDE drives.
HTTrack Website Copier	Downloads the contents of a Web site to a local directory, and recursively builds all directories, getting the HTML code, images, and all other files from the server to your computer.
Hurricane Search (formerly WinGrep)	Displays the results of a search in a hierarchical tree of the files that contain a match.
IE Forensic Tool	Generates an easy-to-read report showing all of the cookies, as well as the Web browsing historyof the current user.
IECacheList (Lite-Free, or Pro - Commercial)	Display the contents of Internet Explorer's index.dat files, including "lost" and hidden content.
IECookiesView	IECookiesView is a small utility that displays the details of all cookies that Internet Explorer stores on your computer.
IEhist	Dumps the Web browsing history from Internet Explorer index.dat files into delimited files suitable for import into other tools.

Table 6.1 Continued

IEHistoryView	Reads all information from a history file and generates the list of all URLs that you have visited in the last few days, and displays it.
Incident Response Collection Report (IRCR)	A collection of tools that gathers or analyzes forensic data on a Windows system.This productis similar to The Coroner's Toolkit (TCT) by Dan Farmer and Wietse Venema. This program is
Index Dat Spy	Exposes the contents of any index.dat file, even if the file is currently in use by Windows. It translates binary data into a human-readable form.
Inforenz Forager®	Analyzes the hidden history of computer files.
Innovision USB WriteBlocker	Examines USB devices in Windows
Inquire	Issues a SCSI Inquiry command and lists any hard disk drives found along with model number, product revision level, and serial number (ESN). This is a Windows application.
Internet Cache Explorer	Displays the content of a browser cache. It comes with many additional features, like a search function that you can use to find URLs or text within visited pages and an option to save complete Web pages (including images and all) for later use.
IrfanView	Displays 32-bit image data. Many supported file formats. Can be usedas an external viewer for forensic analysis.
JAFAT - Archive of Forensic Analysis Tools	Obtains the Safari browsing and download history on a Mac OS X system. Also, the archive contains a link parser, a cookie tool, and "dumpster dive."
Karen's Power Tools	Variety of utilities.
KaZAlyzer	Performs the following operations: lists all database entries in a tabular form, displays the file integrity tag, enablestagging and commenting on each record, identifies files that appear from title, keywords, or other evidence to be child pornography, identifies files that have a known child pornography hash value, identifies all graphics and movie files, sorts by individual columns, exports the content of a database to a CSV file, and produces summary reports. KaZAlyser can open one or more database files from any FastTrack based installation, such as KaZaA, iMesh, and Grokster, and display the contents in a tabular form. Once loaded into KaZAlyser filters can be applied to the database entries to limit the display to particular records such as 'all graphics files" or "identify known child pornography."

Continued

Table 6.1 Continued

	KaZAlyser is the successor to the popular P2Pview KaZaA and Morpheus database viewer. KaZAlyser provides significant enhancements to the forensic investigation process.
KnTTools with KnTList	Acquires physical memory evidence from select Windows operating systems.
List Alternate Data Streams (LADS)	Lists all alternate data streams (ADS) of an NTFS directory. It shows the ADS of encrypted files, even when these files were encrypted with another copy of Windows 2000. It can recursively search through subdirectories and displays a byte total.
LDE	A disk editor for Linux, originally written to help recover deleted files. It is 100-percent free under the Gnu public license.
LibPST	A set of functions in library form for accessing Outlook's Personal Folders. Included with this library is the "readpst" utility that converts a PST file to mbox format. This program is part of the ol2mbox project
Legal Imager and reaSsembly Application (LISA)	LISA is a MS-DOS-based disk-imaging tool, suitable for taking images of hard disk drives for the purpose of forensic analysis.
ListDLLs	Displays the full path names of loaded modules. Also, ListDLLs will flag loaded DLLs that have different version numbers from their corresponding on-disk files (which occurs when the file is updated after a program loads the DLL), and can determine which DLLs were relocated because they are not loaded at their base address.
Live View	Creates a VMware virtual machine out of a raw (dd-style) disk image or physical disk. This tool is Java-based.
Lookout	Performs lightning-fast search for e-mail and files. Works with Microsoft Outlook.
M2CFG USB WriteBlock	Toggles write protection for USB devices on and off.This utility requires Windows XP Service Pack 2
M2CFG Yahoo! Email/Text Parser	Reads a Yahoo! e-mail text file, and writes each e-mail message with its respective attachment(s) to a separate file.Yahoo! returns the contents of each screen name as one text file. This file contains all e-mail content and base64 attachments as plain text.
Mac Emulator	Macintosh emulator for Window XP.
MACMatch	Searches for files by their last write date, last access date, or creation date without changing any of this data.

Table 6.1 Continued

MacOpener 2000	Accesses Macintosh-formatted disks from a Windows PC. Enables access to to Macintosh disks form the Windows desktop. You can also format disks as Macintosh disks.
mac-robber	Collects data from allocated files in a mounted file system. The data can be used by the mactime tool in The Sleuth Kit to make a timeline of file activity.
Magic Rescue	Scans a block device for all recoverable file types and calls an external program to extract them. It looks at the "magic bytes" within the contents of a file, so it can be used both as an undelete utility and for recovering a corrupted drive or partition.
Mailbag Assistant	Mailbag Assistant supports Outlook Express, Eudora, Netscape, Mozilla, Pegasus, The Bat!, Forte Agent, Calypso, PocoMail, FoxMail, Juno 3.x, UNIX mail (Pine, Elm, mbox, etc.), and EML message files.
MailMeter Forensic	E-mail management and investigative solution for organizations that need to perform in-depth forensic analysis and evidentiary discovery on corporate email data
MAK_HTML	Links all files in a folder to an index.htm file that can be used to "browse" the identified files. A program from Dane Mares.
MD5	MD5 Hashing algorithm.
MD5 and Hashing Utilities	Calculates the 128-bit MD5 hash value of any file or block of text.
MD5Deep	Computes MD5 message digests on an arbitrary number of files.
md5summer	Generates and verifies MD5 checksums. md5summer is an application for Microsoft Windows 9x, Windows NT, Windows ME, Windows 2000, and Windows XP.
Media Merge/PC	Reads a tape in any format and also look at any part of the tape in unprocessed mode. In order to do forensic analysis on data from a tape, first it is essential to read the tape. Often with an investigation, tapes may be obtained but no knowledge of how they were written. With Media Merge/PC, the raw data may always be read provided a compatible tape drive is available, and the chances are extremely high that the logical tape format will automatically be detected and the files restored just as on the host system.
memdump	Dumps main memory (/dev/mem) on Solaris, BSD Unix, and Linux systems.

Continued

Table 6.1 Continued

memfetch	Dumps the memory of a running process, either immediately or when an error condition occurs. It recovers information that would otherwise be lost, making it easier to check the internals of a running process.
Metadata Assistant	Analyzes Word, Excel, PowerPoint 97, PowerPoint 2000, PowerPoint 2002 (XP), and PowerPoint 2003 documents to determine what hidden metadata a client might see, display its findings, and then offer the ability to clean the document by selecting a variety of options.
MiTeC Tools	Includes an EXIF reader, Windows Registry recovery, a a portable executable reader, and other utilities.
Mod Com	Alters the operating system files on a floppy boot disk so that, when booted, it will not alter anything on the C: drive. This is what is done manually in the basic forensic classes when you alter boot disks to disable access to the C: drive. This program creates a forensically sound boot disk.
Mount Image Pro	Enables the mounting of EnCase, UNIX DD, or SMART forensic images as a drive letter on your file system.
MS Access Database Cracker	Decrypts the master password stored in a Microsoft Access database. There are two utilities in the zip for decoding Access 95, 97, 2000, and XP.
Mtools	Mtools is a collection of utilities to access MS-DOS disks from UNIX without mounting them. It supports Win-95 style long file names, OS/2 XDF disks and 2m disks (store up to 1992k on a high density 3 1/2 disk).
NASA Tools	Fatback and Enhanced Loopback.
NCFS Software Write-block XP	Blocks write operations to
Nero Express	Burns CDs and DVDs.
NetAnalysis - Forensic Internet History Analyser.	Forensically analyzes Internet user activity.. It can also extract Web browser history from unallocated disk space.
NetIntercept - Network Forensics Analysis Tool	Captures, analyzes, and discovers network traffic. The Forensic View uncoveers malicious activity. It also searches connections by several customizable criteria.
Nigilant32	An incident response tool thatcaptures as much information as possible from a running system with the smallest potential impact.
NirSoft Web site	Provides a unique collection of small and useful freeware forensic analysis utilities.

Table 6.1 Continued

NT SAMs	Linux boot disk, that accesses the Windows Partition then resets account passwords by exploiting that SAM file.
NTLast	Performs security audits on Windows NT systems. NTLast is specifically targeted for security and IIS administration. It performs scheduled reviews of NT event logs to uncover server breaches. Identifies and tracks users who have gained system access and documents the details. . It reports on the status of IIS users and well as filters Web server logons from console logons.
ntpasswd	Edits Registry entries and NT passwords. An image disk to boot in order to recover (reset) password by modifying the encrypted password in the Registry's SAM file. You'll need the rawrite to create the boot disk ou with any UNIX dd if=bootdisk.bin of=/dev/fd0 bs=1024.
ntreg	Mounts NT Registry files such as SAM and SECURITY on Linux file systems. This is a file system driver for Linux that accesses the NT registry file format. Supports only read-only mounts.
Open Digital Evidence Search and Seizure Architecture (ODESSA)	A cross-platform framework for performing computer forensics and incident response.
OE-Mail Recovery	Recovers Outlook Express e-mail.
Offline Registry Parser	A Perl script that parses the raw Registry files in binary mode, and prints out the data, to include last-write information.
Online Digital Forensic Suite (OnlineDFS)	Applies digital forensics technology to the investigation of operational computer systems.
Ontrack PowerControls	Copies and searches mailbox data in Microsoft Exchange Server backups, unmounted databases (EDB files), and information store files.
Open Ports	Detects all open TCP and UDP ports on a computer system, including the owner process. Using advanced port-to-process mapping technology developed for Port Explorer, OpenPorts displays results in five different styles, allowing for easy interpretation by both scripts and human eyes. This is a CLI (command line interface/console) tool.
OST to PST	Creates a PST from an orphaned OST file.

Continued

Table 6.1 Continued

Outindex E-Discovery & Compliance	Extracts emails from Microsoft Outlook PST files and writes them to databases such asOracle and SQL Server.
Outlook Recovery	Recovers data from corrupted Microsoft Outlook Personal Storage files (PST).
Oxygen	Forensically analyzes mobile phones. The software does not change any data on the phone. Oxygen runs on any version of theWindows operating system.
P2 Power Pack	Contains the following items from Paraben Forensics: Case Agent Companion version1.0, Decryption Collection Enterprise version 2.5, E-mail Examiner version 4.01, Forensic Replicator v3.1, Forensic Sorter version 1.0, Network E-mail Examiner version 1.9, PDA Seizure version 3.0.1.35, Text Searcher version 1.0, and Chat Examiner version1.0.
Paraben Case Agent Companion	Optimizes the time of the examiner and the agent working the case. Includes viewers for over 225 file formats and is compatible with Paraben's P2.
Paraben Cell Seizure	Enables the forensic acquisition of user data and portions of unallocated storage on some devices. Cellphone forensics is not to be compared with traditional bit stream forensics. Cellphone data storage is proprietary, based on the manufacturer, model, and system. Each device is unique and should be dealt with caution as each cellphone has unique considerations. Continual advances will be made to Paraben's Cell Seizure in reference to. acquiring of proprietary data. Paraben's Cell Seizure currently supports certain models of Nokia, Sony-Ericcson, Motorola, and Siemens. Paraben's Cell Seizure also supports GSM SIM cards with use of a SIM card reader,which can be found in Cell Seizure Toolbox
Paraben Decryption Collection	A suite of programs that recover passwords. Includes EFS support, Windows Server 2003 support, and Lotus Notes support along with support for everything included in our Standard Edition. It also includes QuickBooks 2003 support and Peachtree 2004 support.
Paraben E-mail Examiner	Examines e-mail messages.. E-mail Examiner claims to recover more active and deleted mail messages than the leading competitor.

Table 6.1 Continued

Paraben Forensic Replicator	Replicates exact copies of drives and media. Paraben's Forensic Replicator can acquire a wide range of electronic media, from a floppies to hard disks. Forensic Replicator images can be compressed and segmented and easily read into the most popular forensic analysis programs.
Paraben Forensic Sorter	Classifies data into over 14 different categories, recovers deleted files, and filters common hashes (for example, FOCH), making examinations easier to manage and faster to process.
Paraben NetAnalysis	Interrogates Web browser cache and history data with powerful searching, filtering, and evidence identification.
Paraben Network Email Examiner	Examines a variety of network e-mail archives. Network E-mail Examiner is compatible with E-mail Examiner and can easily be comfigured for more complex tasks.
Paraben PDA Seizure	Captures data and reports on data from a PDA. Runs in a Windows environment. Includes USB support.
Paraben Text Searcher	Searches text..
Paragon Ext2FS Anywhere	Mounts Linux partitions on Windows file systems as normal logical drives with appropriate drive letters.
Partition Image	Saves partitions in many formats to an image file. Runs on Linux and other versions of UNIX.
Partition Magic	Formats and manages partitions on a hard drive.
Pasco	Examines the contents of Internet Explorer's cache files. Pasco parses the information in an index.dat file and outputs the results in a field-delimited manner so that it may be imported into your favorite spreadsheet program. Pasco runs Windows, Mac OS X, Linux, and BSD UNIX operating systems.
PC Inspector™ File Recovery	Recovers data on FAT 12, FAT16, FAT32, and NTFS file systems. Some of the features in PC INSPECTOR File Recovery 3.x include finding partitions automatically, even if the boot sector or FAT has been erased or damaged This feature does not work with the NTFS file system.
PC Inspector™ Smart Recovery	Recovers data from Flash Card, Smart Media, SONY Memory Stick, IBM Micro Drive, Multimedia Card, Secure Digital Card, or other storage media for digital cameras.

Continued

Table 6.1 Continued

Palm dd (pdd)	Images memory and forensically acquires data from the Palm OS family of PDAs. pdd will preserve the crime scene by obtaining a bit-for-bit image or "snapshot" of the Palm device's memory contents. Runs on all versions of the Windows operating system.
PDU spy	Another mobile phone examination program. The site has a number of interesting recovery programs and useful bits for investigating phones. Freeware.
PenguinBackup	The PalmPilot single-floppy backup system.
PERKEO++	Locates pornography on any type of digital media.
PFC Viewer	Displays and exports the contents of the AOL Filing Cabinet, also known as the Personal Filing Cabinet. PFC Viewer is a Java application.
Phonebase	Reports on the contents of SIM cards and phone memories.These are typically lists of phone numbers and associated names, recently made calls and text messages.
Photo Rescue	Retrieves deleted dataon Compact Flash cards.
PhotoRec	Recovers lost pictures from digital camera memory.
Picalo	Forensically analyzes data. Data Analysis and Fraud Detection, a collaborative, open-source effort to produce a data analysis application suitable for auditors, fraud examiners, data miners, and other data analysts. Windows, Macintosh, and UNIX/Linux versions.
Pictuate	Finds and analyzes images for pornographic content; scans computer disks to find all image files, then scores those images as to their content. The results are sorted in descending order and viewed in a panel of thumbnail images. The user sees the images most likely to be pornographic (a target image) first and progresses to less likely candidates in subsequent panels.
Pilot-Link	Retrieves the contents of ROM and RAM from Palm devices. Additionally pilot-xfer allows acquisition
Pinpoint Labs Free Tools	Pinpoint FileMatch, Pinpoint Hash, Pinpoint Metaviewer.

Table 6.1 Continued

Port Explorer	Displays all open ports on a system and the programs that own them — also called port-to-process mapping. Additional functionality includes packet sniffing, bandwidth throttling, and country detection. Port Explorer has an intuitive GUI that allows displays all network activity on your computer.
POSE	Emulates Palm devices on the desktop.
Private Eyec	Displays the entire contents of an area on your PC referred to as *protected storage*.
Process Dumper & Memory Parser (MMP)	Creates a dump of a running process for forensic purposes. MMP parses the metainformation stored within process dumps made with Process Dumper (pd) and extracts the different process mappings to disk.
Process Explorer	Identifies the handles and DLLs that have been opened by processes or loaded into process space. The Process Explorer display consists of two subwindows. The top window always shows a list of the currently active processes, including the names of their owning accounts, whereas the information displayed in the bottom window depends on the mode that Process Explorer is in. If it is in handle mode you'll see the handles that the process selected in the top window has opened; if Process Explorer is in DLL mode you'll see the DLLs and memory-mapped files that the process has loaded. Process Explorer also has a powerful search capability that will quickly show you which processes have handles opened or DLLs loaded. The unique capabilities of Process Explorer make it useful for tracking down DLL-version problems or handle leaks, and provide insight into the way Windows and applications work.
Process Viewer for Windows (PrcView)	Displays detailed information about processes running under Windows.
procshow	Displays information about a running process.
Protected Storage Explorer	Displays saved data from the Protected Storage Service, including passwords for e-mail accounts in Microsoft Outlook, Microsoft Outlook Express, MSN Messenger, saved Internet Explorer form data such as phone numbers, credit card numbers, web email, search engine queries, user names and passwords on Web pages, and cached logon credentials of sites that require authentication, including FTP sites.

Continued

Table 6.1 Continued

Password Recovery Toolkit (PRTK)	Recovers passwords. Developed by Access Data.
PTfinder Collection	A collection of PTfinders for Windows 2000, Windows XP, Windows XP SP1, Windows XP SP2. and Windows Server 2003.
PurgeIE	This versatile program allows you either to manage your cookies and cache files in a logical manner or, for complete privacy, to eliminate them altogether, complete with all references, tracks, trails and strays.
Putty	PuTTY is a free implementation of Telnet and SSH for Win32 and UNIX platforms, along with an xterm terminal emulator.
PyFlag	PyFlag is the Python implementation of Flag - a complete rewrite of FLAG in the much more robust python programming language. Many additional improvements were made.
QuickTime	Plays video, audio, QuickTimeVR, and graphics files. Freeware.
Remote Data Acquisition (RDA)	Remotely acquires data — such as disk cloning or disk/partition imaging data — and verifies the transfer using MD5 or CRC32 checksums. The program is both the server and the client and is a command-line tool that runs on Linux.
R-Drive Image	Creates disk image files for backup or duplication purposes. The image file contains an exact, byte-by-byte copy of a hard drive, partition, or logical disk.
RealPlayer	Plays every major media format, including DVDs.
Recover	Automates some steps of recovering deleted files on an /Ext2 filesystem.
Recover CHK Files	Renames CHK files to the correct extension.
Recover4all	After downloading the utility, run the self extracting EXE file and choose the *unzip* option. You should not save anything to the drive where your deleted files are to prevent the deleted files from being overwritten. For example, if your deleted files are on the C: drive, do not install any software and do not download anything to C:. Download Recover4all to A:. Recover4all does not require installation and can be run from a floppy disk.
Red Cliff Web Historian	Reviews Web sites stored in the history files of the most commonly used browsers, including: Microsoft Internet Explorer, Mozilla, Firefox, Netscape, Opera, and Safari.

Table 6.1 Continued

Regdat & RegdatXP	Displays the contents of copies of the Win9x and WindowsMe Registry files System.dat, Classes.dat, and User.dat, and displays the contents in a Regedit like interface. You can search for keys and values and export them. Also functions to compare the file with the current Registry are provided as well as tools to edit the file.
Registry Information Extractor (RIE)	Extracts Registry information on Windows 95, Windows 98, Windows ME operating system. It will be updated to extract a lot more information from the Registry, including Windows NT, Windows 2000, and Windows XP support. Currently It can extract the following information: Registered Owner, Registered Organization, Windows Version, Windows Version Number, Windows Installed Date & the Computer Name. RIE can also be used as a File Viewer from within EnCase. This is a test release of a software utility that is in development and under testing.
RegMon	Determines the applications that are accessing the Windows Registry, which keys they are accessing, and the Registry data they are reading and writing – all in real-time.
RegTools	Contains the following applications: RegDACL - , a permissions manager for Registry keys; RegOwner - ownership manager for Registry keys. RegAudit - audit manager for Registry keys. RegLast - lists or queries the last write data of Registry keys.
Resource Hacker	Views, modifies, renames, adds, deletes, and extracts resources in 32-bit Windows executables and resource files (RES files). Freeware.
Resplendent Registrar	Reviews the Registry in Windows NT, Windows 2000, Win9x, and Windows ME systems. It searches the Registry extremely fast, and presents the search results in a usable fashion.
Restorer2000	Undeletes files deleted accidentally in NTFS partitions, and recovers data on formatted or corrupted drives.
Rifiuti	Examine the contents of the INFO2 file in the Recycle Bin, and produces field-delimited output that can be imported into your favorite spreadsheet program. Rifiuti runs on Windows, Mac OS X, Linux, and BSD Unix platforms.
RKDetect	Finds services hidden by generic Windows rootkits such as Hacker Defender.

Continued

Table 6.1 Continued

R-Linux	Recovers data on Ext2 file system partitions used in several UNIX operating systems. Supported host operating systems: Windows 9x, Windows Me, /Windows NT, Windows 2000, Windows XP. Recovered data can be written to any disk visible by the host operating sytem. R-Linux also can create disk images that can be later processed by more powerful R-Studio. Freeware.
R-Mail	Recovers accidentally deleted e-mail messages and recovery damaged DBX files where Outlook Express stores folders with e-mail messages. The new e-mail data recovery technology IntelligentRebuild allows R-Mail users to quickly reconstruct damaged DBX files created by Outlook Express and easily restore the lost messages. The messages are recovered in EML format and can be imported into Outlook Express e-mail and Internegt news databases.
RootkitRevealer	Detects rootkits. Processes files produced by Windows NT 4 and higher.
R-Studio	Recovers and undeletes data. Supports the FAT 12, FAT 16, FAT 32, NTFS, NTFS 5 , Ext2 file systems and recovers files both on local logical and physical disks and disks on remote computers over the networks, even if their partition structures are damaged or deleted.
R-Undelete	Undeletes data on the FAT, NTFS, NTFS 5, and Ext2 file systems. R-Undelete can undelete files on any valid logical disks accessible by the host OS. It cannot undelete files on damaged or deleted volumes, or if the hard drive has been repartitioned.
R-Word	Recovers MS Word documents.
SafeBack	Creates mirror-image (bitstream) backup copies of hard disks or partitions.
SagePass	Retrieves passwords from Sage Sterling, Line 100, and Instant accounting systems.
Scalpel: A Frugal, High Performance File Carver	Reads a database of header and footer definitions and extracts matching files from a set of image files or raw device files.
sdd	A replacement for a the “dd” utility.
SecCheck	Detects and removes malicious applications, back doors, Trojans, worms, and viruses.

Table 6.1 Continued

SecReport - Click 'Toolbox'	Collects security-related information from Windows-based system (SecReport) and compares any two reports from different systems or the same system after a specified delta of time. A suite of two command-line tools.
SecretExplorer	Views, analyzes edits, imports, and exports protected storage , retrieve AutoComplete passwords and data, recovers passwords; transfer passwords and settings from one computer to another.
Secure Hash Signature Generator	Creates signatures that are unique to the data stored on a disk drive. These signatures are used to verify intentional or accidental tampering with the drive image.
SHA verify	Calculates the MD5 (128-bit), SHA1 (160-bit), SHA2 (256-bit), SHA2 (384-bit), and SHA2 (512-bit) hashes of files.
ShoWin	Displays information about Windows such as account passwords. ShoWin displays information a cursor is dragged over it. One of the most popular uses of this program is to display hidden password editbox fields (text behind the asterisks). This will work in many programs although Microsoft has changed the way things work in some of their applications, most notably the MS Office suite and Windows 2000. ShoWin will not work in these cases. Neither will it work for password entry boxes on Web pages, at least with most Web browsers. Additional features include the ability to enable windows that have been disabled, unhide hidden windows , and force a window to stay on top or be placed below others.
SIM Manager	Recovers phone numbers and SMS messages from a range of mobile phones
SIM Scan	Investigates the contents of SIM cards. Freeware.
SIMCon	Provides forensic imaging and analysis of the contents of SIM cards, including recovery of deleted items. Free for law enforcement.
Sleuthkit Windows Executables	Microsoft Windows executables for The Sleuth Kit. Full source code and documentation can be downloaded.
Sleuthkit/Autopsy Foremost patch	Integrates Foremost into Autopsy and enables the configuration file to be edited. Created by P. Bakker.
Sleuthkit/Autopsy Searchtools patch	Performs indexed searching capabilities for Sleuthkit/ Autopsy tools. Created by P. Bakker.

Continued

Table 6.1 Continued

SMART	Monitors hard disk drives and performs failure prediction operations. It uses S.M.A.R.T. technology to monitor the health status of hard disk drives, prevents data loss, and predicts possible drive failure.
SnapView HTML Viewer	Examines recovered HTML pages stored in unallocated space. This viewer is built on the same technology as used by Internet Explorer and loads pages quickly. It can also toggle between page and source view. It can use any available Internet Explorer plug-ins , giving it a large selection of supported file formats. Some of the supported file formats are : HTML, JPEG, GIF, ICO, Flash Move, Adobe Acrobat, Word, Excel, PowerPoint, Bitmap, PNG, and ART.
SNIFFER	Extracts files from the free space of computer hard disks, from a series of disk image files or directly from the hard disk device. Run on MS-DOS.
SomarSoft Utilities	Contains the following programs: DumpSec - a security auditing program for WindowsNT, and Windows 2000, DumpEvt - a Windows NT program to dump the event log in a format suitable for importing into a database, and DumpReg - a program for Windows NT and Windows 95 that dumps the contents of the Registry.
Spam CSI	Performs forensic analysis of e-mail data.
Spector CNE	Records everything employees do online, including instant messaging, chats, sending and receiving e-mails , visiting Web sites, launching applications, downloading files, and typing keys.
Spider	Displays all of the URLs and cookies stored in the index.dat file, and removes them.
ssdeep	Creates a short text signature for each input file. The signatures can be used to match other files against the original. Unlike MD5 or SHA-1, however, this algorithm can match two input files even if they are not exactly the same. Files match if they have significant homologies, or the same sequences of bytes in the same order. Functions in a similar manner to md5deep.

Table 6.1 Continued

Stealer	Extracts the machine name, usernames, and dial-up user accounts and passwords. It also identifies passwords and usernames on secure web sites and password-protected shared folders on a network. Most of this information is stored within the PWL file. Stealer must be run on a restored drive if you are using it to identify information on a seized computer. One law enforcement agency used it to gain access to encrypted data as the password for the encrypted material had been duplicated. This utility can save you weeks of waiting if you are contemplating a brute force attack. Runs on Windows 9x and Windows Me.
StegAlyzerAS	Performs digital forensic analysis. This utility is designed to extend the scope of traditional digital forensic examinations by allowing the examiner to scan suspect media for artifacts of steganography applications.
StegDetect	Automatically detects steganographic content in image files. It is capable of detecting several different steganographic methods for embedding information in JPEG images. Supported methods are: jsteg, jphide (UNIX and Windows), invisible secrets, and outguess 01.3b.
StegHide	Embeds a secret message in a cover file by replacing some of the least significant bits of the cover file with parts of the secret message. After that, the secret message is imperceptible and can only be extracted with the correct pass phrase. Features include support for JPEG, BMP, WAV, and AU files; encryption of plain data before embedding (blowfish encryption algorithm); pseudo-random distribution of hidden bits in stego file; and embedding a CRC32 checksum in the data.
Stego Suite 4.1	Investigates, detects, analyzes, and recovers of digital steganography. Stego Suite 4.1 includes Stego Watch, an automated steganography investigation scanning software package; nine steganography detection algorithms covering all common digital image file types and audio WAV files; Stego Analyst, a visual image analysis package for in-depth digital image and audio file analysis; and Stego Break, an automated steganography cracking tool.

Continued

Table 6.1 Continued

Stellar Phoenix FAT	Recovers data on FAT file systems. It does not to be installed prior to a data loss. Stellar Phoenix FAR provides easy recovery after the partition table, boot sector, and root directory have been corrupted and data has been lost. This software runs on MS-DOS, Windows 95, Windows 98, Windows Me, Windows NT (FAT file systems), Windows 2000 (FAT file systems). Supports all variants of FAT file system.
Stellar Phoenix Novell	Recovers vital data from corrupt volumes resulting from a crashed server. It accesses data when the volumes cannot be mounted. This software can save recovered data to a disk drive after cases of corrupt or missing volumes, deleted files, allocation errors, and partition loss.
Stellar Phoenix NTFS	Recovers data after disk crashes caused by accidental formatting, viruses , software malfunctions, file or directory deletion, or sabotage. Stellar Phoenix NT Data Recovery software examines an inaccessible hard drive and displays the stored data. The disk recovery process requires only a selection of the target files and directories to copy to a working drive.
Sterilize	Sterilizes media for use as working/examination data copies.
STG Cache Audit	Displays the contents of browswer caches, cookie, and history information.
STrace for NT	Examining the system calls made by a process running on Windows NT. It is meant to be used like the strace (or truss) UNIX operating system variants. Runs on Windows NT.
Stream Explorer	Reveals alternate file streams on NTFS file systems. Runs on Windows NT 4.0, Windows 2000, Windows XP, and Windows 2003.
Streams	Reveals NTFS alternate streams.
Strings	Searches for ANSI and Unicode strings in binary images.
SuperDIR	Lists all files and directories on FAT file systems, and dumps the results into a database. It will calculate a CRC32 for every file. Very fast.
SuperSCAB	Analyzes seized computers.
Swapper	Swaps all byte pairs in a file. This is required for forensic research of disks, mobile phones, and copiers.
Symantec	AV program

Table 6.1 Continued

TAFT - The ATA Forensics Tool	Communicates directly with the ATA controller to retrieve information about a hard disk, and to examine and change HPA and DCO settings.
TaFWeb Whois	A version of the "whois" utility.. It can report IP numbers as well as domain names.
TapeCat	Forensically analyzes tapes. Runs on all versions of Windows.
TASK	Performs a comprehensive forensic analysis of Windows and UNIX file systems. Open source from @stake.
Tcpflow	Captures data transmitted as part of TCP connections (flows), and stores it in a way that is convenient for protocol analysis or debugging. The "Tcpdump"utility shows a summary of the packets transmitted on the wire but usually doesn't store the data. In contrast, Tcpflow reconstructs the actual data streams and stores each flow in a separate file for later analysis. Tcpflow processes sequence numbers and correctly reconstructs data streams regardless of retransmissions or out-of-order delivery. However, it currently does not process IP fragments — flows containing IP fragments will not be recorded properly
Tcptrace	Analyzes TCP dump files. It takes as input the files produced by several popular packet-capture programs, including Tcpdump, snoop, etherpeek, HP Net Metrix, and WinDump. Tcptrace can produce several different types of output containing information on each connection seen, such as elapsed time, bytes and segments sent and received, retransmissions, round trip times, window advertisements, and throughput. It can also produce a number of graphs for further analysis.
TCTUtils	An enhancement to the The Coroner's Toolkit (TCT) utility. TCTUtils can list directory inode contents, find the inode that is using a given block, and enables the viewing of inode and block details in various formats.
TestDisk	Checks and undeletes partitions.
Text2Hex	This utility will convert ASCII characters to hexadecimal values. This is particularly useful when performing searches using software that can accept hexadecimal values as search criteria.
ThumbsDisplay	Displays all thumbnail images with original file names and timestamps.
ThumbsPlus	Displays thumbnail images and displays, edits, and batch converts regular image files.

Continued

Table 6.1 Continued

TNEF	Unpacks Microsoft MS-TNEF MIME attachment files.
Tree Browser (TB) / KDE Tree Browser (KTB)	Displays file hierarchies.
True Time	Prompts the user for the correct date and time, and obtains the system date and time from the system BIOS. The output can be redirected to a file for use in forensic investigations. Excellent addition to a forensic boot disk.
TULP2G	Extracts and decoding data stored in electronic devices. Based on Windows .NET technology.
UDP Cast	A Netcat substitute.
UndeleteSMS	Recovers deleted SMS messages from a GSM SIM card.
Unique	Eliminates duplicate records in a file. This program will take a sorted input file and copy records to the output for which it finds a unique occurrence of the sort key. The program passes the input file, and when it finds a new and unique sort key in a record it copies that record to the output, and disregards all subsequent records that contain that same sort key. Therefore, only a single record per sort key is copied to the output file.
URLSearch	Extracts web addresses and other text strings from local and remote files. Results can be edited and exported in various ways. Web addresses can be sorted by domains and exported to your browser's bookmarks (supported browsers are Netscape, Internet Explorer, and Opera). UrlSearch can also be used as a download agent and can view history and cookie files, edit typed URLs, and calculate hit rates.
USB Blocker	This application was designed to meet the needs of write blocking USB devices.
UTF-8 output patch for task-1.60/Sleuth Kit-1.6x	Displays foreign characters in the names of files locateds in an NTFS image. By default, Sleuth Kit converts the Unicode NTFS file system structures into ASCII. This patch changes that behavior, but doesn't change the keyword searching ability. That is still done with the "strings" and "grep" utilities, which do not easily support Unicode.
Ultimate Toolkit (UTK)	The complete AccessData software kit.
VFC - Virtual Forensic Computing	Rapidly boots a forensic image of a suspects computer, or boots a write-blocked physical hard drive.
Viesoft Forensic Scan	Scans "index.dat" files.

Table 6.1 Continued

vinetto	Extracts thumbnail images and associated metadata from the "thumbs.db" files generated by Windows. A console-based program that runs under Cygwin or Linux.
Vision	Reports all open TCP and UDP ports and maps them to the process or application accessing them.
VNC	Remotely accesses and views.
Web Cache Illuminator	Creates a meaningful list that displays all file names in the Web browser cache and the title of the Web page.
WebDate	Returns the last modified date and time of a Web site, Web page, or individual file.
Webtracer	Determinesthe owner of a Web site, the location of a server, the sender of an e-mail, and other evidence of Internet identity.
WhatFormat	Analyzes the first bytes of a file for signatures (or magic numbers)and determines the format of the file from this evidence.
Win32 Analyzer	Creates a forensic snapshot of a computer. Uses Windows and third-party tools.
Win32 First Responder's Analyzer Tookit	Performs basic security tasks. A self-extracting Windows executable designed to be run from a floppy disk.
Windows Forensic Toolchest (WFT)	Provides automated incident response on a Windows system and collects security-relevant information from the system.
Windows Forensics and Incident Recovery tools	A set of tools produced by the Forensic Server Project for incident response and computer forensic activities on Windows systems.
Windows Memory Forensic Toolkit (WMFT)	A collection of proof-of-concept tools, used for forensic analysis of physical memory images gathered from compromised Windows machines
Windows NT/2000 Incident Response Tools	Gathers and analyzes forensic data on a Windows system and produces snapshots of past states of system. This program is a collection of tools, and like The Coroner's Toolkit (TCT), most of the tools are oriented towards data collection rather than analysis.

Continued

Table 6.1 Continued

WinDump	Creates dump files of network traffic data. WinDump prints the headers of packets on a network interface that match a specified boolean expression. It can also be run with the "-w" flag, which causes it to save the packet data to a file for later analysis, andwith the "-b" flag, which causes it to read from a saved packet file rather than to read packets from a network interface. In all cases, only packets that match the specified expression will be processed by WinDump.
WinGREP	Searches files for specified text strings.
WinHex	Edits hexadecimal data.
WinImage	Creates disk images from a floppy and vice versa, among other functions. Developed because Windows lacks a native "dd" command.
Wininterrogate	Recursively scans a directory structure and obtains the following information: file name and size, complete path of files and directories, directory name, , creation time, last access time, last write time, and MD5 checksum. Additional information gathered from DLL, VBX, DRV, EXE, OCX, BIN, SCR files includes CompanyName, FileDescription, FileVersion, InternalName, LegalCopyright, OriginalFilename, ProductName, ProductVersion, Comments, LegalTrademarks, PrivateBuild, and SpecialBuild.
WinOra	Converts the installation time format used in the Windows Registry.
WinPcap 3.1	Used in several network investigation tools.
WinRAR	Compresses and uncompresses RAR-format files. Also processes other compression formats. Runs on all versions of Windows.
WinZip	Compresses and uncompresses ZIP-format files. Runs on all versions of Windows.
Wipe	Securely wipes files from storage media. It is based on work by Peter Gutmann.
X1 Search	Searches for information, previews it in its native format, and takes action on it regardless of type or location.
XANALYS Forensic Analyzer	Automatically examines an EnCase file. Extends EnCase functionality.
XnView	Views and browses multimedia files, and converts the format.
X-Ways Trace	Tracks and examines pastWweb browsing activity on a computer.

Table 6.1 Continued

Yahoo Message Archive Decoder	Accesses Yahoo Messenger archive files (DAT files) and presents them in a readable format.
ZefrJPG	Recovers JPG files destroyed by the Love Letter worm and its variants on FAT or NTFS file systems (Thanks to Robert Green @ http://personal.atl.bellsouth.net/~lasrpro.)
Zeitline	Edits timelines for computer forensic purposes.r
Zero Assumption Digital Image Recovery	Recovers digital image data produced by a digital camera. digital images. It allows you to recover digital photos accidentally deleted from digital camera memory. Freeware.

Part 2: Forensic Hardware Tools

The features and advantages of hardware computer forensic tools are based on techniques that substitute software for processes and the use of equipment that is compact and easy to operate.

These devices are often powered from the source or from the suspect machine. However, the forensic analyst who relies on forensic hardware must ensure that all possible connectors are available prior to starting a job. Some of the advantages include:

1. The mbedded development that has been completed, saving space and time and generally simplifying the acquisition process.
2. The greater portability of the products.
3. The increeased speed of acquiring of digital data using hardware devices compared to using software.

Hard Disk Write Protection Tools

Hardware protection devices offer a simple method to acquire an image of a suspect drive with much less fiddling with the configuration settings in software. This makes the process simpler and less prone to error.

The following section introduces a number of these hardware-based forensic tools.

NoWrite

NoWrite prevents data from being written to the hard disk. It supports hard disk drives with high capacities. It is compatible with all kinds of devices including USB or FireWire boxes, adapters, and IDE interface cables. It supports communication between common IDE interfaces.

NoWrite is only functional on native IDE devices. It supports all USB features such as plug- and-play. It is compatible with most operating systems and drive formats. NoWrite is transparent to the operating system and the application programs — MS-DOS, Windows, Linux, Sun, and Macintosh operating systems are incapable of detecting it. NoWrite maintains the

image of the evidence obtained from a system and the data is able to be viewed using EnCase, FTK, and Ilook in preview modes. This product is created by MyKey Technologies, Inc.

FireWire DriveDock

Wiebetech's FireWire DriveDock is a forensic instrument designed to load hard drives on computer systems. It is a 3.5-inch hard drive that is used along with a single device to give complete FireWire desktop storage. It is a compact device of about four cubic inches that can control everything in a 3.5-inch hard drive.

At the base of the drive is a mounting plate where the power and FireWire cables are attached and an AC adapter. The drive volume does not require additional drivers. It may be used in all Macintoshes and PCs with a FireWire port. The data transfer rate is a minimum of 35 MB per second.

LockDown

Lockdown by Paraben is an advanced Firewire-to IDE or USB-to-IDE write blocker that combines speed and portability in enabling IDE media to be acquired quickly and safely in Windows-based systems. Paraben's Lockdown System includes the following equipment:

- Paraben's Lockdown
- 40-wire IDE ribbon cable for legacy desktop hard drives
- 80-wire IDE ribbon cable for newer desktop hard drives
- 44-pin-IDE-notebook-to-80-pin-IDE-desktop cable converter
- 4-pin (mini) to 6-pin FireWire (IEEE 1394a) cable
- 6-pin FireWire cable
- USB 1.0/2.0 cable
- Molex power cables for desktop hard drives
- External power supply
- User uanual
- Hard carrying case

The devices feature the following capabilities:

- Connections using FireWire or USB interfaces
- Hot-swapping of suspect hard drives without restarting the computer
- Small size (4&" wide, 3" deep, and 1" high) for portability and ease-of-use in the field
- IDE ports for "desktop IDE" and "laptop IDE" media, negating the need for a desktop-to-laptop IDE adapter

- Allows the reading of IDE hard drives through lightweight, portable, and common computer port interfaces (Firewire or USB)
- Can be used to acquire drives through Windows using a process that is substantially faster than MS-DOS-based acquisitions
- Easy to hook up and start an analysis

Write Protect Card Reader

The Write Protect Card Reader from Intelligent Computer Solutions transfers data from digital cameras, digital camcorders, PDAs, MP3 players, and digital voice recorders to a laptop or desktop computer. It can read several types of flash memory while blocking any writes to it. It is a small palm-size package with a simple USB 1.1/2.0 connection and requires no external power.

The features of the Write Protect Card Reader include:

- High-speed USB 2.0 connection
- Backward-compatible to USB 1.1
- The ability to read 12 types of digital media in a plug-and-play manner: CF-I, CF-II, Smart Media, Memory Stick™, Memory Stick PRO, Memory Stick DUO, Memory Stick PRO DUO, Micro Drive, Multimedia Card, Secure Digital Card, MINI Secure Digital Card, XDcard (using optional CF adapter)
- Can read data from four different media types simultaneously
- Maximum data throughput is 480 Mbits/sec using a USB 2.0 connection
- Unique icon for each media type in the Windows "My Computer" folder
- Small enough to fit in a jacket pocket
- Bus powered [WTW1] no AC adapter needed

Drive Lock IDE

The DriveLock IDE hard drive write protection from Intelligent Computer Solutions (ICS) prevents write commands from being accidentally sent to hard disk drives connected through the IDE or PATA hard drive interfaces. Itdevice also blocks Serial ATA hard drives using the SATA option and blocks write commands sent to the hard drive while previewed or duplicated.

Serial-ATA DriveLock Kit

The DriveLock Serial-ATA device also fromIntelligent Computer Solutions is a hardware write-protection device designed to prevent data writes to SATA, IDE, and PATA hard disk drives. The tool is connected to a computer's PATA interface to block write commands sent to the hard drive while the drive is being previewed or duplicated.

Wipe MASSter

Wipe MASSter from ICS is a commercial drive wiper. It has the following features:

- Wipe up to nine hard drives at speeds exceeding 2.6 GB/minute for one pass.
- Conforms to U.S. Dept. of Defense Standard DoD 5220.22-M
- Forensic audit printout after wipe
- Virtually no degradation in speed when adding drives
- Works with any IDE, EIDE, or Ultra-DMA hard drives.
- Supports 3.3-volt hard drives
- Easy firmware upgrading
- Software upgrading and saved log files using PCMCIA CompactFlash memory cards
- Automatic cable test before every run
- Locks unit setting with supervisor password
- Stores customized, predefined setting configurations
- 48-bit support for drives greater than 137 GB
- Unlocs and wipes HPA/DCO areas
- Mix different size and type of drives during the same load

ImageMASSter Solo-3 IT

Designed exclusively for forensic data acquisition, the ImageMASSter Solo-III forensics data imaging tool from ICS is a lightweight, portable handheld device that can acquire data to one or two evidence drives in high speed, with a speed exceeding 3 GB/minute.

Image MASSter Solo-3 Forensic is a complete, inclusive high-speed data acquisition tool that integrates all of the latest advanced features in data imaging. It is capable of capturing data from IDE and laptop drives, Serial ATA and SCSI drives, and Flash Cards.

With the integrated Link MASSter Option also unopened computers through both FireWire and USB2 ports.

Features of the Image MASSter SOLO 3 Forensics are the following:

- MD5 and CRC32 Hashing: Simultaneous MD5 and CRC32 hashing generation can be performed during the data capture.
- Touchscreen User Interface: The unit's advanced Touchscreen user interface and programmed keyboard provides for ease of use.
- High Speed Operation: Data acquired with a transfer rate that exceeds 3 GB/minute.

- Built in Write Protection.
- Built in FireWire 1384B and USB 2.0 Interface: Capture data from an unopened notebook or PC. Safely preview the drive while connected to the ImageMASSter Solo III write-protected port.
- Capture to Two Evidence Drives Simultaneously: Capture data to two hard drives simultaneously at high speeds, without speed reduction.
- Multiple Capture Methods: Acquire data using forensic sector-by-sector format or segmented file format methods (as does the "dd" utility in Linux) which enables copying multiple images to one evidence drive.
- WipeOut: Sanitize drives and erase data at speeds greater than 3 GB/minute.
- Biometric Security: This option uses fingerprint recognition to ensure that only qualified and authorized users are allowed to operate the ImageMASSter Solo III forensics unit.
- Audit Trail and Logs: Detailed operational event log information can be printed or saved to CompactFlash media.
- Multiple Media Support: Supports data transfer between P-ATA, S-ATA, and SCSI hard disk drives. Also supports creating images of data on Flash Media devices and notebook computer drives.
- Upgradeable: Software and firmware upgrades performed using the supplied CompactFlash card.

ImageMASSter 4002i

The Intelligent Computer Solutions Image MASSter 4002 loads two IDE hard drives at speeds exceeding 1.94 GB/minute. It has the following features:

- Prints a forensic audit report after the wipe operation.
- Virtually no degradation in speed from adding drives.
- Works with any IDE, EIDE, or Ultra-DMA hard drives.
- Mix sizes and models of drives during the same load operation.
- Loads any operating system and application software, including Windows 95, Windows 98, Windows NT, Windows 2000, Windows XP, UNIX, Linux, and MacOS.
- Full verification and error reporting: zero-defect copies.
- Automatic load size detection.
- Built-in options include NTFS-IQcopy and Wipeout.
- Supports 3.3-volt hard drives electrical interface.

- Easy Firmware upgrads.
- Software upgrading and saving log files to PCMCIA CompactFlash memory cards.
- Automatic cable test before every run.
- 48-bit addressing for drives greater than 137 GB.
- Ability to lock a unit setting with a supervisor password.
- Ability to store multiple, customized, and predefined setting configurations.
- Block-map copy mode included.
- The Image MASSter 4002 Hard Drive Duplicator includes a database option.

ImageMASSter 3002SCSI

Intelligent Computer Solutions' ImageMASSter 3002SCSI is a High Speed disk imager provides operating speeds that exceed 3.9 GB/minute. No speed degradation occurs when copying to four hard drives. It offers the capability tocopy data to four target drives simultaneously with no speed degradation. It has the following features:

- Multiple Media Support: The ImageMASSter3002SCSI SCSI duplicator supports current SCSI interfaces such as Ultra-160 and Ultra-320, including SCSI Narrow (50-pin), SCSI Wide (68-pin) and SCSI-SCA80 (80-pin). Supports data transfer between P-ATA and SCSI hard disk drives using optional SCSI to P-ATA converters. Also supports SCSI to S-ATA using optional SCSI to S-ATA converters.
- Multiple Copy Methods: Copy data using a sector-by-sector method, which supports copying any operating system and file system. Copy only allocated data using the ICS proprietary intelligent copy method known as "IQCOPY", which supports Windows FAT-16, FAT-32, and NTFS file systems.
- Partitions and Formats Drives: Target drives are automatically partitioned and formatted during the copy process, eliminating the need to manually prepare a drive prior to usage.
- Copy between Different Drive Models and Sizes: Simultaneously copy data between different drive models and capacities. Partitions will automatically be scaled to the size of the target drive.WipeOut: Sanitize (per DoD specs) up to five SCSI drives simultaneously at speeds greater than 3.5 GB/minute.
- Bad Sector Handling: Provides flexibility in handling bad sectors found on the source drive.
- Log Information: Stores detailed operational event log information.

Image MASSter 3004 SATA

The Image MASSter 3004 SATA from ICS can duplicate four Serial ATA (SATA) hard drives simultaneously. It has the following specifications and features:

- Transfer rate can exceed 1.6 GB/minute depending on drive models in use. Transfer rate of 1.6 GB/minute demonstrated in a laboratory environment using Seagate ST360015AS and ST3120023AS hard drives series.
- Interface: Serial ATA data connector and Serial ATA power connector.
- Mix sizes and models of drives during the same load operation.
- Support for image copy and IQCOPY mode.
- Support for FAT16, FAT32, and NTFS partitions using IQCOPY mode. Support for all other file systems using image copy mode.
- Windows XP-based user interface.
- Image MASSter 3004 SATA Hard Drive Duplicator includes built-in Wipeout option.
- Built-in MultiMASSter option.
- 48-bit support for hard drives with capacities greater than 137 GB.
- Logs hard drives under Windows for more flexibility.
- Mixes any combination of regular IDE and Serial ATA hard drives in the master and target position (which enables copying the contents of regular IDE to Serial ATA hard drives and vice versa)

Table 6.2 lists several forensic hardware tools.

Table 6.2 Forensic Hardware Devices

Hardware Device	Description
NoWrite	NoWrite prevents data from being written to the hard disk. It supports hard disk drives with high capacities. It is compatible with all kinds of devices including USB or FireWire boxes, adapters, and cables belonging to IDE. It supports communication between common IDE interfaces.
FireWire DriveDock	FireWire DriveDock is a forensic instrument designed to load hard drives on computer systems. It comprises of a 3.5-inch hard drive that is used along with a single device to give complete FireWire desktop storage. It is a compact device of about 4 cubic inches that would control everything in a 3.5-inch hard drive.

Continued

Table 6.2 Continued

Hardware Device	Description
LockDown	Lockdown by Paraben is an advanced Firewire or USB to IDE write-blocker that combines swiftness and portability to allow IDE media to be acquired quickly and safely in Windows based systems.
Write Protect Card Reader	The Write Protect Card Reader transfers data to a computer system from digital cameras, digital camcorders, PDAs, MP3 players and digital voice recorders. It can read multiple types of flash memory while blocking any writes to it. It is a small palm-size package with a simple USB 2.0/1.1 connection and requires no external power.
Drive Lock IDE	The DriveLock IDE Hard Drive Write Protection is designed to completely prevent write commands from being accidentally sent to hard disk drives connected through the IDE or PATA hard drive interfaces. This write protect device also blocks Serial ATA hard drives using the SATA option. It is designed to block write commands sent to the hard drive while previewed or duplicated.
Serial-ATA DriveLock Kit	The DriveLock Serial-ATA device is a hardware write protect device designed to prevent data writes to SATA, IDE and PATA hard disk drives. The tool is connected to a computer's PATA interface in order to block write commands sent to the hard drive while being previewed or duplicated.
Wipe MASSter	Wipe MASSter is a commercial drive wiper.
ImageMASSter Solo-3 IT	Designed exclusively for Forensic data acquisition, the ImageMASSter Solo-III Forensics data imaging tool is a light weight, portable hand-held device that can acquire data to one or two evidence drives in high speed, exceeding 3 GB/Min.
ImageMASSter 4002i	The Image MASSter 4002 loads 2 IDE Hard drives at speeds exceeding 1.94 GB/Min.
ImageMasster 3002SCSI	ImageMasster 3002SCSI is a High Speed disk imager with operating speeds that can exceed 3.9 GB/min. No speed degradation occurs when copying to 4 hard drives.
Image MASSter 3004SATA	The Image MASSter 3004 SATA can duplicate 4 Serial ATA hard drives simultaneously

Summary

The primary purpose of this chapter is to gain a level of familiarity with the wide range of hardware and software forensics tools that are available. It is not possible to detail of all the features and functions of the tools in this chapter, therefore, the student should investigate these features and functions on their own.

The two primary areas that will be examined are:

- Software-based tools
- Hardware tools and write blockers

Frequently Asked Questions

Q: There are so many tools. Do we need them all?

A: It is important to become familiar with a wide range of tools. It is not necessary to have all the tools.

Q: Many of the tools run on an operating system I am not familiar with. Do I need to become familiar with these operating systems?

A: Although it is good to have some level of familiarity with a wide range of systems, these tools are available for most operating systems. Whether you are familiar with Linux, Windows, or another operating system, it is likely that you will be able to find a tool that does the function you need that runs on that system.

Q: What types of evidence can digital forensics tools provide?

A: Computers store large amounts of data to a network or system of hard disks. Much of this information is stored without the user being conscious of its existence. This data may be in the form of tangible files or information that the computer used to carry out a specific task. A few examples are user files, system files, deleted files, and system data that enable the computer to perform its tasks.

Q: If I am unsure what has occurred and think that the case may never go to court. Do I still need to image the drive?

A: You never know what may occur. Imaging the drive is useful if something does turn out to be wrong. Also, being able to demonstrate that you use imaging tools as a standard practice will make it more difficult to your procedures to be attacked in court, even on other unrelated cases.

Q: Do I need to scan a drive for malware if I have a read-only image?

A: Any forensic image file with a capture containing 100 percent of the information on a hard drive has the potential to be a virus pathway into a secure computer system. Any forensic investigation must exercise appropriate care in ensuring that none of these pathways exist.

Chapter 7

Incident Response: Live Forensics and Investigations

Solutions in this chapter:

- **Postmortem versus Live Forensics**
- **Today's Live Methods**
- **Case Study: Live versus Postmortem**
- **Computer Analysis for the Hacker Defender Program**
- **Network Analysis**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

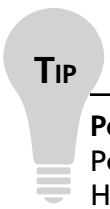
Introduction

To pull or not to pull the plug, that is the question. Today, cyber crime investigators are faced with the grueling task of deciding whether shutting down a computer system is the most efficient and effective method to gather potential electronic evidence. Traditionally, computer forensics experts agreed that shutting the computer system down in order to preserve evidence and eliminate the potential changing of information is best practice prior to examination. I remember having the phrases "shut it down," and "don't change anything" beaten into my brain during the numerous trainings I've attended throughout the years. However, one of the fundamental misconceptions with this philosophy is that computer forensics is the same as physical forensics. I would argue that they are not the same, given that computer forensics technology changes faster than traditional forensics disciplines like ballistics, serology, and fingerprint analysis. The second misconception is that we always collect everything at a physical crime scene. In a physical forensics environment, we commonly photograph the physical crime scene and take "reasonable" precautions to ensure the evidence is not disturbed. The truth is, in many cases, we only collect samples from a physical crime scene.

Nevertheless, we have accepted this methodology as best practice, and have backed ourselves into a litigation corner. The evolution of technology has put us face to face with the harsh reality that it is sometimes more advantageous to perform "Live" analysis than a "Postmortem" one. The problem is that live analysis often changes evidence by writing to the hard drive. File time stamps, Registry keys, swap files, and memory are just some of the items that can be affected when conducting analysis on a live computer system. Often, once the live analyst is done, the resulting MD5 hash will not match the hash collected prior to the live collection.

Postmortmem versus Live Forensics

Why should we even consider conducting live investigations as a valid forensic methodology? The reason is we have to! In the pages that follow, I will discuss the need to move away from traditional methods of computer forensics toward a live forensics model.



TIP

Postmortem vs. Live Forensics
Postmortem and live forensics are both great evidence gathering techniques. However, in cases where you can only conduct a postmortem forensics, the need to look at other systems within the environment is strengthened. This expansion of your scope to include other systems on the network will give you a better understanding of how the target system acted within its native environment.

Evolution of the Enterprise

Technology has evolved in such a way that conducting live investigations is really the only option you have under certain circumstances. In the days of old, computer networks were simple. In today's world, the evolution of the enterprise network work makes it difficult for system administrators, IT security personal, and the like to be at more than one location. Managing IT resources at a single site can be a daunting task. Now think of the larger corporate network schema. Many companies have multiple computers at a single location. Additionally, those corporations may also have several locations in a city, country, or continent. What would happen to our resources if we had to respond to every site and pull the computer off the network to conduct a forensic analysis for every suspected compliance issue, security breach, or compromised host? This would be even worse if after all the effort, time, and resources, we conclude that none of the aforementioned even occurred. Sound familiar? It should, because it happens every day in the cyber world. Triage is a common practice when diagnosing problems within a network. It is our first reaction, and we don't necessarily assume we are under attack, or that our systems have been compromised. In a live forensic environment, IT security personnel could log on remotely, view running processes, dump physical memory, and make an educated guess as to whether or not the computer should be imaged remotely, or be physically removed from the network for further analysis. In this scenario, the investigator, using live forensics techniques, doesn't have to physically respond to the location to address the issue until they are satisfied with their initial inquiry. This methodology will help conserve resources.

Evolution of Storage

Now back to pulling the pull. Once upon a time there was a server. This server was about 630 terabytes (TB) in size. It was responsible for handling the day-to-day operations of Company X, which traded stocks for its clients 24 hours a day. This server was believed to be compromised because of some unusual traffic detected within the log files of the firewall. This scenario presents us with the following issues. Problem 1: How are we going to fit this 630TB image into our 250 GB USB2 external drive? Problem 2: How long would it take to image a drive that size? Problem 3: The machine can not be shut down because the company would suffer a financial loss. Let's discuss the preceding problems one at a time.

Problem 1: It's not possible. You will need a bigger drive.

Problem 2: The data resides on a substantially large server (630TB). Imaging the entire server is not practical, even though best practices dictate we should. Here is one of the reasons why: 630TB is equal to 6,926,923,254,988,880 bytes. 630 × 1,099,511,627,776 (1 Terabyte) = 6,926,923,254,988,880 bytes. See Table 7.1 to determine the byte sizes used in this scenario.

Table 7.1 Byte Conversion Chart

Drive Size	Numerical Representation	2 to the Following Power
1 kilobyte	1,024	10
1 megabyte	1,048,576	20
1 gigabyte	1,073,741,824	30
1 terabyte	1,099,511,627,776	40
1 petabyte	1,125,899,906,842,624	50
1 exabyte	1,152,921,504,606,840,000	60

Let's assume you use the ICS Image MASSter Solo-3 IT, which states it can duplicate hard drives at a rate of 3 GB a minute.

- Divide 6926923254988880 / 3221225472 (3 gigabytes) = 2150400 total minutes
- Divide 2150400 minutes /60 minutes (1 hour) = 35840 total hours
- Divide 35840 total hours / 24 hours (1 day) = 1493 total days
- Divide 1493 days / 365 day (1 year) = over 4 years to image the entire drive.

As you can see from the preceding bullets, imaging the entire one-to-one drive is not practical. Even if you imaged the data, by utilizing additional resources, the analysis of such a large volume could prove just as prohibitive. The difference in conducting an analysis on such a large volume, as compared to specific data objects and/or smaller storage systems, (using a detective's analogy) would be equivalent to interviewing every person who lives on a block where a homicide has occurred (reasonable), versus interviewing everyone who lives in the city of the homicide victim (not reasonable).

Notes from the Underground

Using Compression

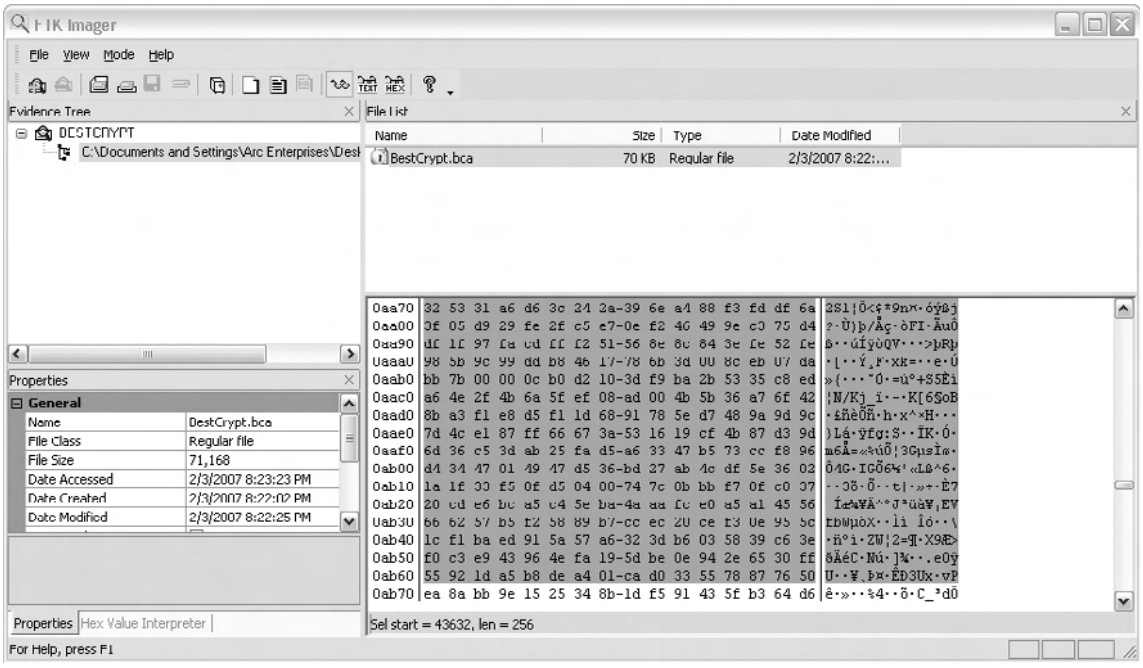
If you're thinking that the use of compression could solve the preceding problems, you would be mistaken. Compression increases the time it takes to image the server's hard drive because the compression algorithm needs to examine and remove the redundant items prior to compressing them. Additionally, it would still be impossible to compress the larger hard drive into the smaller USB external drive.

Problem 3: Shutting down the server is also not an option since the most obvious side effect would be the economic harm Company X would experience as a result. Many systems in existence today are mission critical, such as those supporting health care, transportation, and so on, and they couldn't be shut down without causing detrimental effects.

Encrypted File Systems

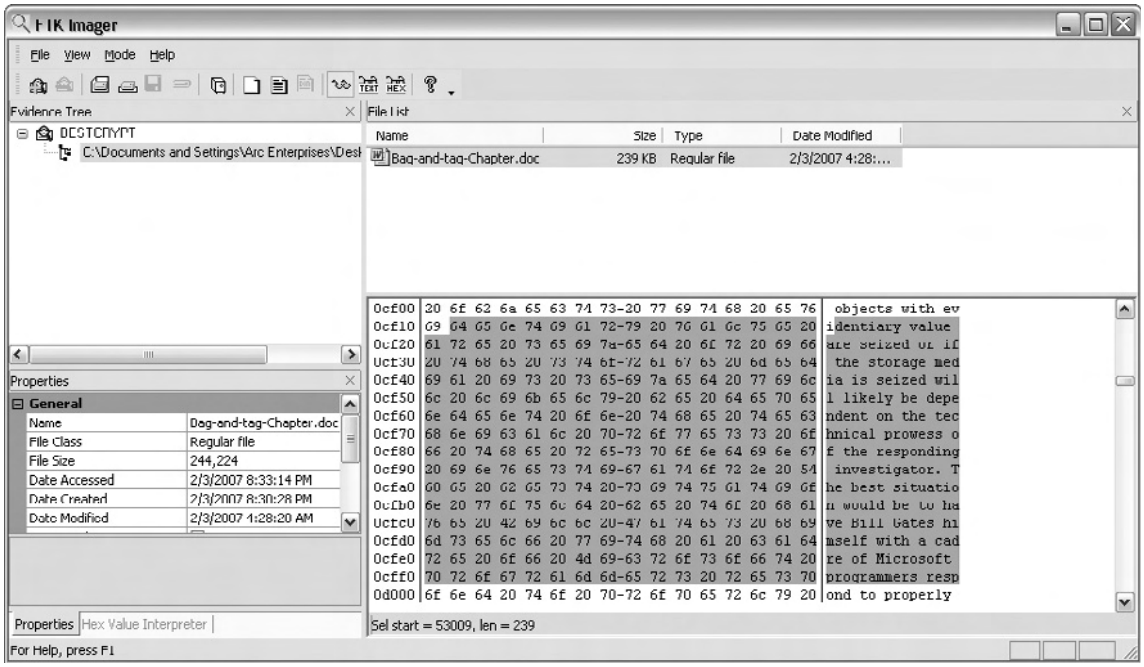
The use of encryption has increased during the last few years. Its increased use presents a unique problem to investigators when conducting postmortem analysis. When encryption is applied to a data object, the contents of that object are illegible. Encryption, by default, is designed to obfuscate, and sometimes compress, the contents of the data object it encrypts. Once encrypted, the object's contents are hidden and are pretty much impossible to interpret. Encryption is applied to these data objects in one of three ways. The first implementation is file level encryption, in which individual files are encrypted. Figure 7.1 shows the contents of an encrypted file.

Figure 7.1 File Contents When the File Is Encrypted Using AccessData's FTK Imager



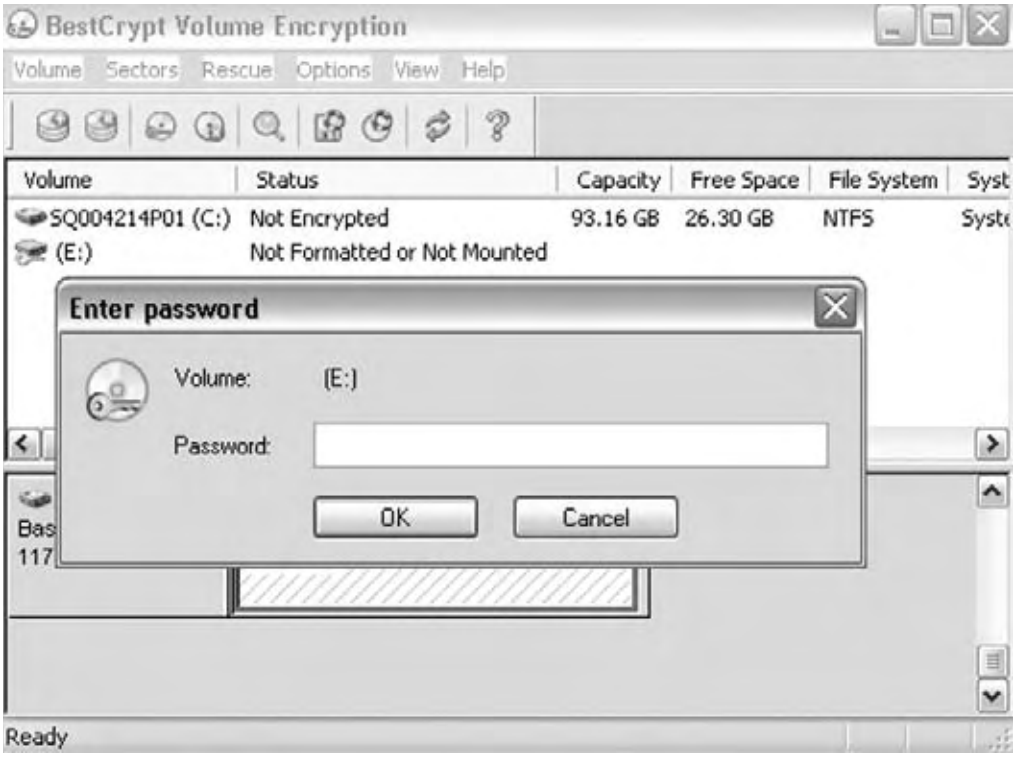
In order for an examiner to perform a postmortem analysis, he must first decrypt the file. Figure 7.2 shows a decrypted file. This could prove extremely difficult if the investigator does not have access to the encrypted file's password. No password may result in having to use a password cracking program. This decrypting process may prove useless if the password is too large, or the file is encrypted with a strong encryption algorithm and implementation.

Figure 7.2 File Contents When the File Is Not Encrypted Using AccessData's FTK Imager


FTK Imager
File View Mode Help
Evidence Tree
DESTCRYPT
C:\Documents and Settings\Arc Enterprises\Desk
File List
Name
Size
Type
Date Modified
Baq-and-tag-Chapter.doc
239 KB
Regular file
2/3/2007 4:28:...
Properties
General
Name
Dag-and-tag-Chapter.doc
File Class
Regular file
File Size
244,224
Date Accessed
2/3/2007 8:33:14 PM
Date Created
2/3/2007 8:30:28 PM
Date Modified
2/3/2007 1:28:20 AM
Properties
Hex Value Interpreter
0cf00 20 6f 62 6a 65 63 74 73-20 77 69 74 68 20 65 76 objects with ev
0cf10 69 64 65 6e 74 69 61 72-79 20 76 61 6c 75 65 20 identiary value
0cf20 61 72 65 20 73 65 69 7a-65 64 20 6f 72 20 69 66 are seized or if
0cf30 20 74 68 65 20 73 74 6f-72 61 67 65 20 6d 65 64 the storage med
0cf40 69 61 20 69 73 20 73 65-69 7a 65 64 20 77 69 6c ia is seized wil
0cf50 6c 20 6c 69 6b 65 6c 79-20 62 65 20 64 65 70 65 l likely be depe
0cf60 6e 64 65 6e 74 20 6f 6e-20 74 68 65 20 74 65 63 ndent on the tec
0cf70 68 6e 69 63 61 6c 20 70-72 6f 77 65 73 73 20 6f hnical prowess o
0cf80 66 20 74 68 65 20 72 65-73 70 6f 6e 64 69 6e 67 f the responding
0cf90 20 69 6e 76 65 73 74 69-67 61 74 6f 72 2e 20 54 investigator. T
0cfa0 68 65 20 62 65 73 74 20-73 69 74 75 61 74 69 6f he best situatio
0cfb0 6e 20 77 6f 75 6c 64 20-62 65 20 74 6f 20 68 61 n would be to ha
0cfc0 76 65 20 42 69 6c 6c 20-47 61 74 65 73 20 68 69 ve Bill Gates hi
0cfd0 6d 73 65 6c 66 20 77 69-74 68 20 61 20 63 61 64 mself with a cad
0cfe0 72 65 20 6f 66 20 4d 69-63 72 6f 73 6f 66 74 20 re of Microsoft
0cff0 70 72 6f 67 72 61 6d 6d-65 72 73 20 72 65 73 70 programmers resp
0d000 6f 6e 64 20 74 6f 20 70-72 6f 70 65 72 6c 79 20 ond to properly
Sel start = 53009, len = 239
For Help, press F1

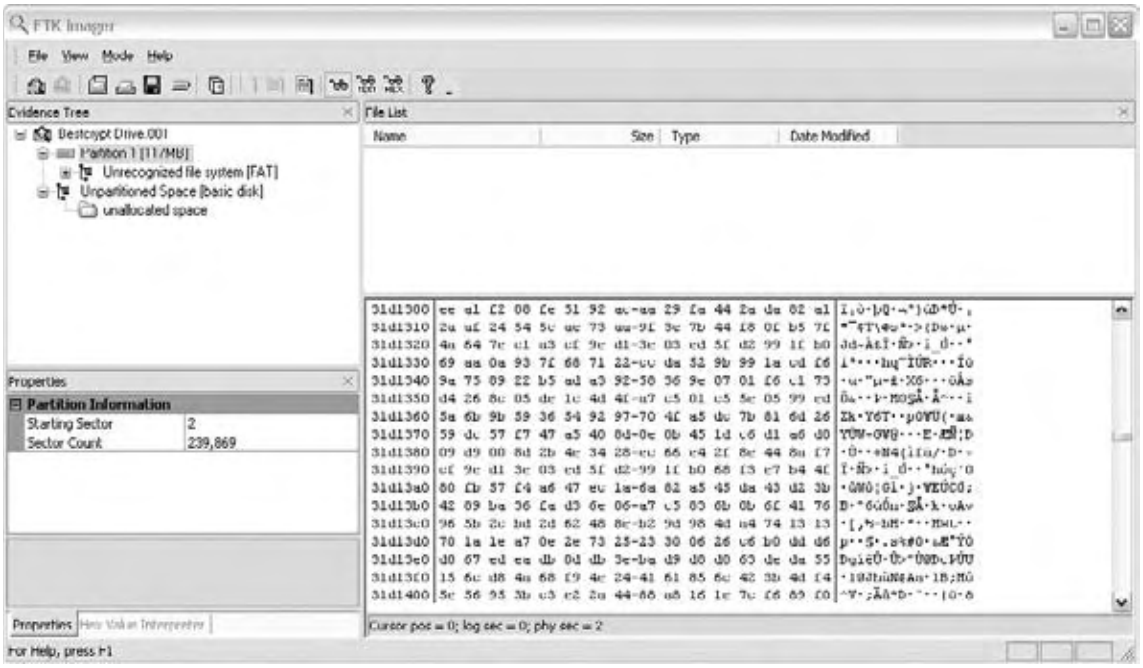
The second method used when applying encryption is volume level encryption. In this case, a volume within the hard disk is encrypted. Figure 7.3 shows an encrypted volume.

Figure 7.3 A BestCrypt Encrypted Volume



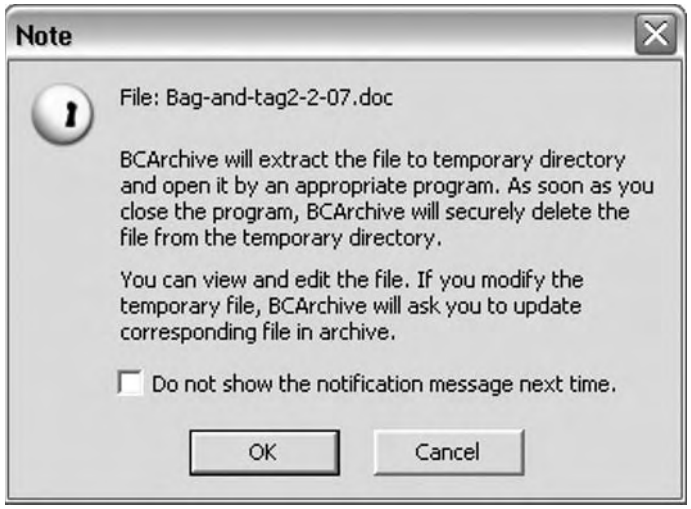
The third method used when encrypting a data object is whole disk encryption. This is when the entire hard drive is in encrypted. Figure 7.4 offers a forensic image of a fully encrypted disk. As you can see, its contents are illegible, and are of little value to a forensic examiner.

Figure 7.4 A Forensic Image of an Encrypted Hard Drive Using AccessData's FTK Imager



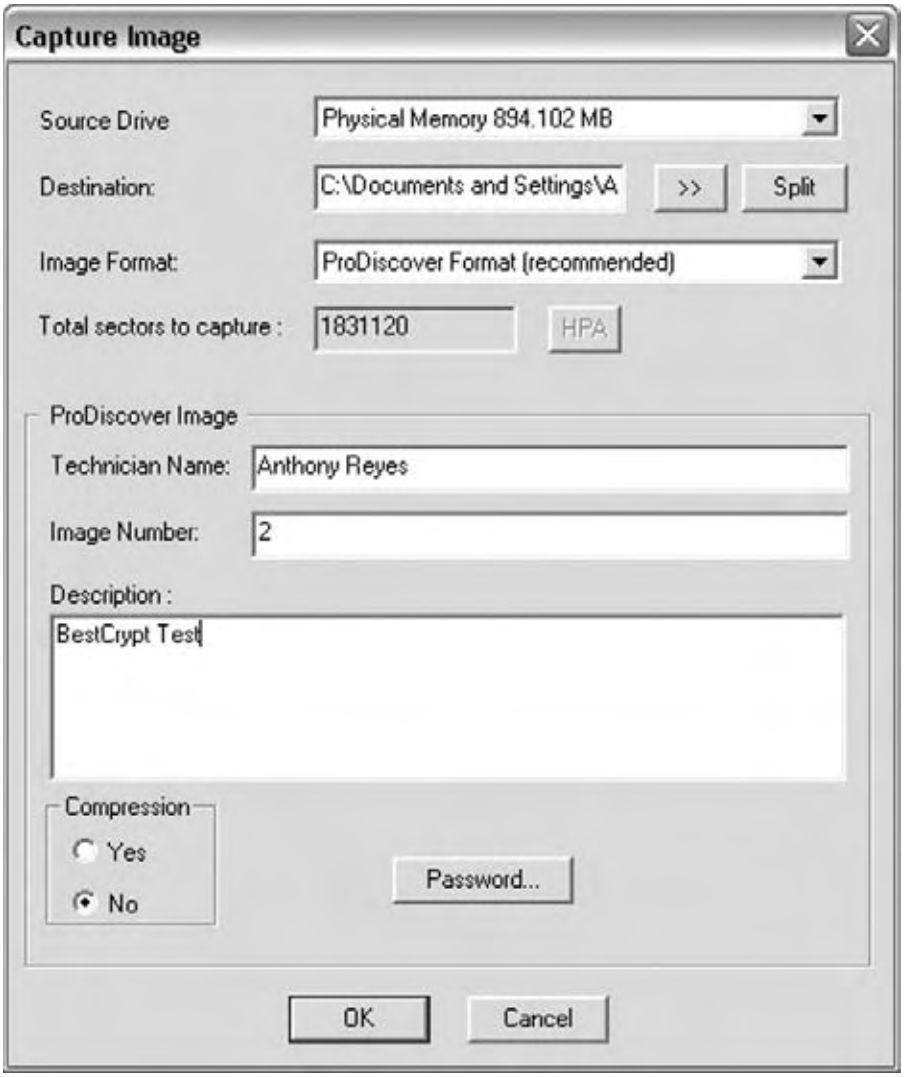
When conducting postmortem forensic analysis against the first two methods, investigators often hope to find artifacts of an encrypted file in its decrypted state that may be left in allocated or unallocated space. These artifacts are sometimes created once the document has been opened, or when the plug has been pulled while the file is still displayed on the screen. While this is a valid premise, recovery of these artifacts may not always be successful. Moreover, performing a proper shutdown may further decrease your chances of finding such evidence. In Figure 7.5, you will notice that the program BestCrypt offers to open the file in a temporary folder, and then securely delete the file when the program is closed.

Figure 7.5 A File Cleaning Operation Offered by BestCrypt


Note
File: Bag-and-tag2-2-07.doc
BCArchive will extract the file to temporary directory and open it by an appropriate program. As soon as you close the program, BCArchive will securely delete the file from the temporary directory.
You can view and edit the file. If you modify the temporary file, BCArchive will ask you to update corresponding file in archive.
Do not show the notification message next time.
OK
Cancel

Using live forensics, the chances are significantly greater to view the contents of the encrypted file. If the document is open, it will most likely be loaded into physical memory. In a live forensic environment, the investigator could image the physical memory of the computer system and glean useful information about what files and programs the suspect may be currently using. So, before pulling the plug, it may be worth our while to examine the contents of the physical memory. Figure 7.6 shows one example of how we could image physical memory by using a network forensics tool.

Figure 7.6 The Technologies Pathways' ProDiscover IR Imaging Screen



Once the image has been created, we can examine its contents. In Figure 7.7, you will notice the contents of the encrypted file are displayed in a readable format in the lower right-hand pane. Recovery of this information is because the file has been unencrypted by the user who is currently working with the document. Additionally, in Figure 7.8 you can see the BestCrypt program is running in physical memory. This information is also displayed in the lower right-hand pane.

Figure 7.7 An Unencrypted Document in Memory Using Technologies Pathways' ProDiscover IR

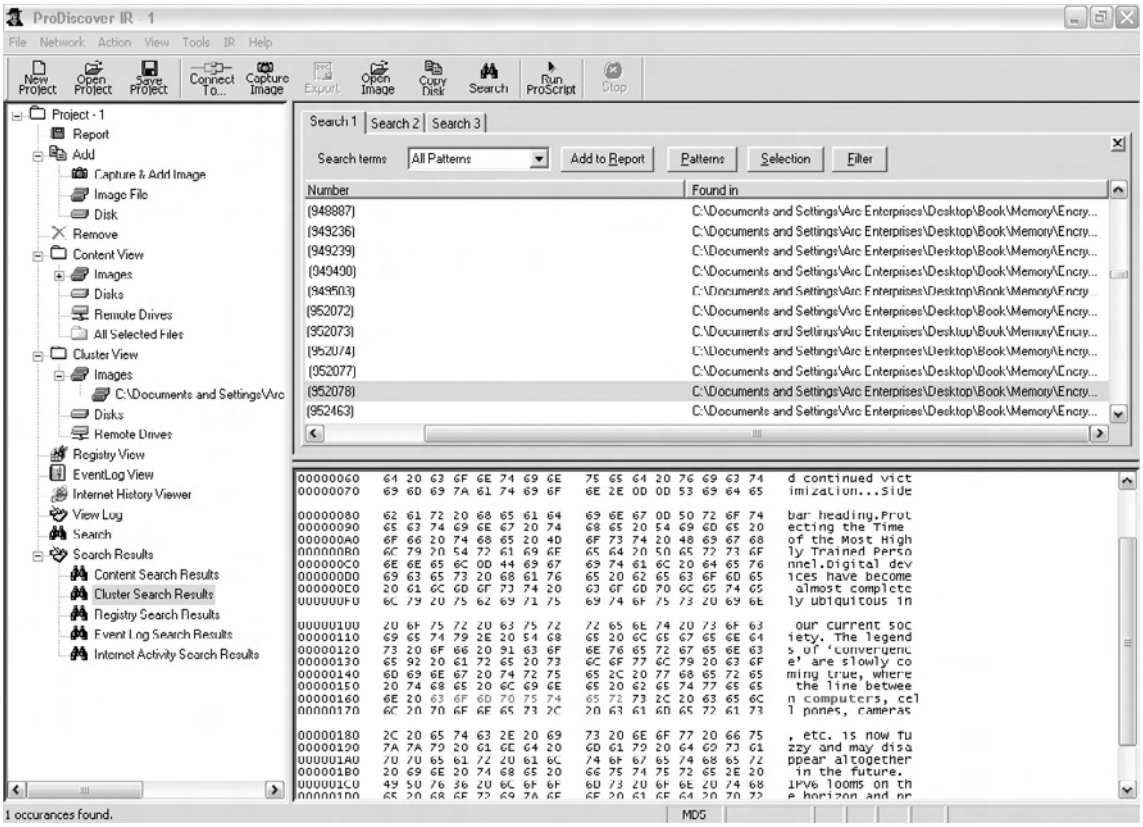
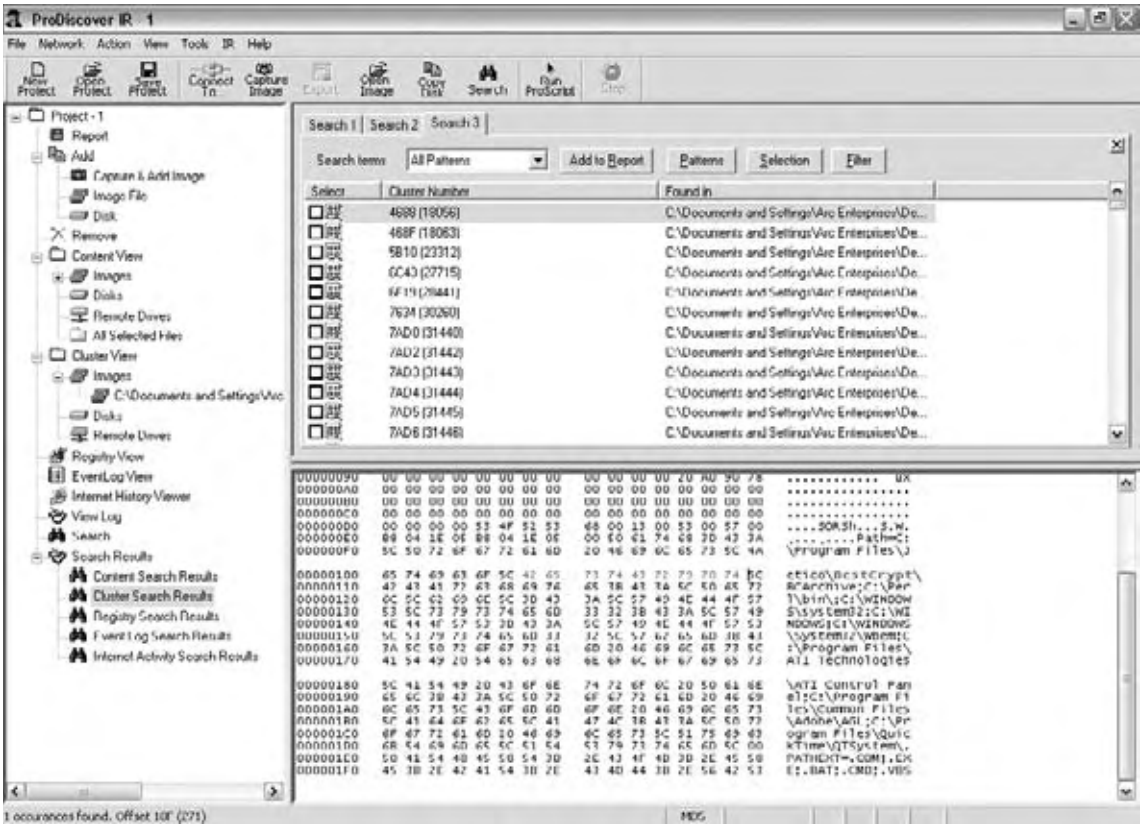
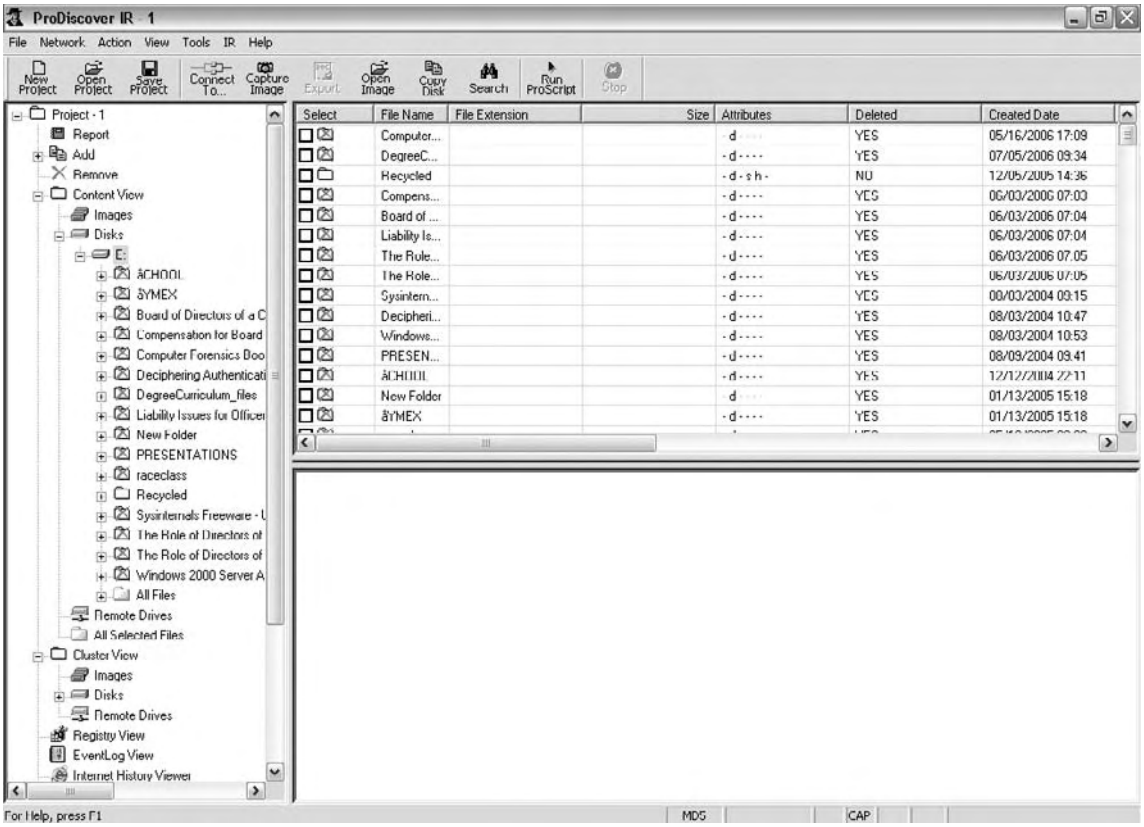

ProDiscover IR - 1
File Network Action View Tools IR Help
New Project
Open Project
Save Project
Connect To...
Capture Image
Export
Open Image
Copy Disk
Search
Run ProScript
Stop
Project - 1
Report
Add
Capture & Add Image
Image File
Disk
Remove
Content View
Images
Disks
Remote Drives
All Selected Files
Cluster View
Images
C:\Documents and Settings\Arc
Disks
Remote Drives
Registry View
EventLog View
Internet History Viewer
View Log
Search
Search Results
Content Search Results
Cluster Search Results
Registry Search Results
Event Log Search Results
Internet Activity Search Results
Search 1
Search 2
Search 3
Search terms
All Patterns
Add to Report
Patterns
Selection
Filter
Number
Found in
(948887)
(949236)
(949239)
(949490)
(949503)
(952072)
(952073)
(952074)
(952077)
(952078)
(952463)
C:\Documents and Settings\Arc Enterprises\Desktop\Book\Memory\Encry...
00000060 64 20 63 6F 6E 74 69 6E 75 65 64 20 76 69 63 74 d continued vict
00000070 69 6D 69 7A 61 74 69 6F 6E 2E 0D 0D 53 69 64 65 imization...Side
00000080 62 61 72 20 68 65 61 64 69 6E 67 0D 50 72 6F 74 bar Heading.Prot
00000090 65 63 74 69 6E 67 20 74 68 65 20 54 69 6D 65 20 ecting the Time
000000A0 6F 66 20 74 68 65 20 4D 6F 73 74 20 48 69 67 68 of the Most High
000000B0 6C 79 20 54 72 61 69 6E 65 64 20 50 65 72 73 6F ly Trained Perso
000000C0 6E 6E 65 6C 0D 44 69 67 69 74 61 6C 20 64 65 76 nnel.Digital dev
000000D0 69 63 65 73 20 68 61 76 65 20 62 65 63 6F 6D 65 ices have become
000000E0 20 61 6C 6D 6F 73 74 20 63 6F 6D 70 6C 65 74 65 almost complete
000000F0 6C 79 20 75 62 69 71 75 69 74 6F 75 73 20 69 6E ly ubiquitous in
00000100 20 6F 75 72 20 63 75 72 72 65 6E 74 20 73 6F 63 our current soc
00000110 69 65 74 79 2E 20 54 68 65 20 6C 65 67 65 6E 64 iety. The legend
00000120 73 20 6F 66 20 91 63 6F 6E 76 65 72 67 65 6E 63 s of 'convergenc
00000130 65 92 20 61 72 65 20 73 6C 6F 77 6C 79 20 63 6F e' are slowly co
00000140 6D 69 6E 67 20 74 72 75 65 2C 20 77 68 65 72 65 ming true, where
00000150 20 74 68 65 20 6C 69 6E 65 20 62 65 74 77 65 65 the line betwee
00000160 6E 20 63 6F 6D 70 75 74 65 72 73 2C 20 63 65 6C n computers, cel
00000170 6C 20 70 6F 6E 65 73 2C 20 63 61 6D 65 72 61 73 l pones, cameras
00000180 2C 20 65 74 63 2E 20 69 73 20 6E 6F 77 20 66 75 , etc. is now fu
00000190 7A 7A 79 20 61 6E 64 20 6D 61 79 20 64 69 73 61 zzy and may disa
000001A0 70 70 65 61 72 20 61 6C 74 6F 67 65 74 68 65 72 ppear altogether
000001B0 20 69 6E 20 74 68 65 20 66 75 74 75 72 65 2E 20 in the future.
000001C0 49 54 76 36 20 6C 6F 6F 6D 73 20 6F 6E 20 74 68 IPv6 looms on th
000001D0 65 20 68 6F 72 69 7A 6F 6E 20 61 6E 20 72 e horizon an re
1 occurances found.
MD5

Figure 7.8 A View of Physical Memory Contents Using Technologies Pathways' ProDiscover IR. Note that the BestCrypt Process Is Running.



In the case of whole disk encryption, a forensic examiner using live forensics techniques would be able to view the content of the drive when it is mounted by the suspect. Simply put, because the drive is presently being used, it is unencrypted. Figure 7.9 demonstrates our ability to view the mounted drive's contents in its unencrypted state.

Figure 7.9 An Encrypted Hard Drive's Contents When Mounted Live with a Forensics Tool Like Technologies Pathways' ProDiscover



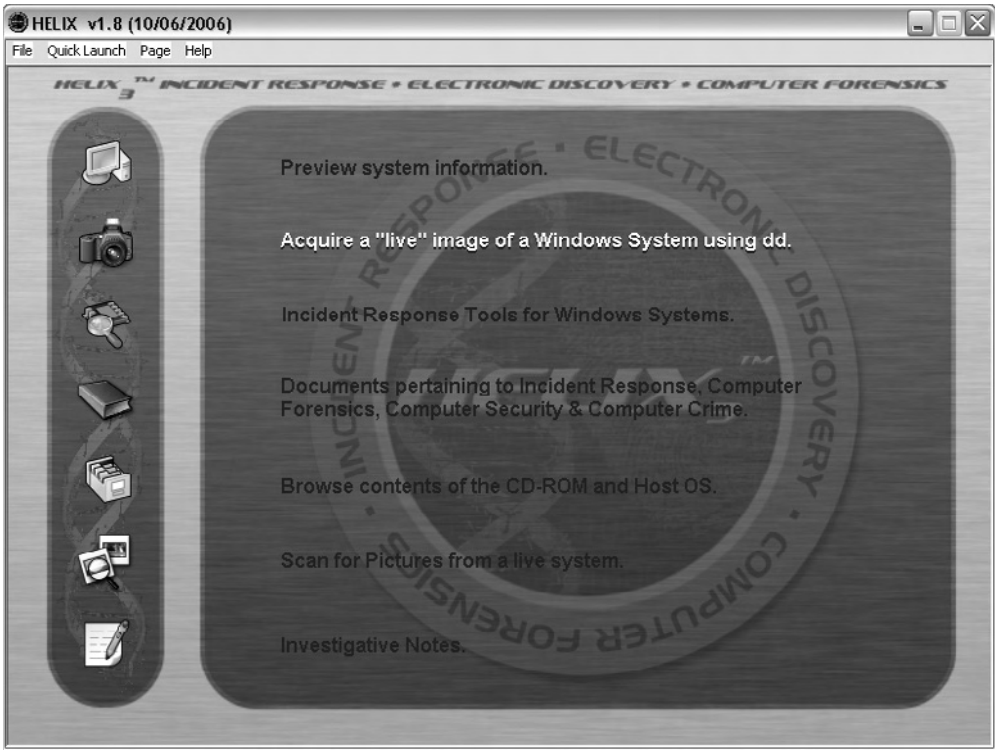
As you can see from the preceding examples, encryption presents a variety of problems for the traditional forensics examiner. With live investigative techniques, however, we can overcome these problems and obstacles.

Today's Live Methods

Several software companies presently manufacture network forensic and investigative software. Guidance Software, Technologies Pathways, Wetstone Technologies, ASR Data, E-fense, and E Trust by CA are just some of the companies that produce this forensic and incident response software. These manufacturers use a variety of methods to conduct live investigations. The first method employed is the Pre-Deployed Agent model, where special software is pre-installed on a computer system prior to an incident. It is usually hidden from the end user and is invoked once it is connected to remotely. The second method currently in use is the Direct Connect model. In this model, the target computer is directly connected to by a remote machine and the software is pushed into memory. The connection remains active until the remote machine is disconnected.

A third method is the On Demand Connection model, where the computer connects to the target machine and pushes the software into memory for a specific task. Once the task issued by the remote machine is completed, the connection is immediately torn down. Finally, some software developers use a boot disk or an investigative CD-ROM. During a live analysis, a disk is loaded to the live machine and a virtual session is initiated with a set of examination tools. Figure 7.10 shows a boot disk that allows you to conduct live forensics, as well as investigations.

Figure 7.10 Is a Picture of the E-fense's HELIX Incident Response, Electronic Discovery, and Computer Forensics Boot Disk


HELIX v1.8 (10/06/2006)
File Quick Launch Page Help
HELIX 3 INCIDENT RESPONSE • ELECTRONIC DISCOVERY • COMPUTER FORENSICS
Preview system information.
Acquire a "live" image of a Windows System using dd.
Incident Response Tools for Windows Systems.
Documents pertaining to Incident Response, Computer Forensics, Computer Security & Computer Crime.
Browse contents of the CD-ROM and Host OS.
Scan for Pictures from a live system.
Investigative Notes.

Case Study: Live vs. Postmortem

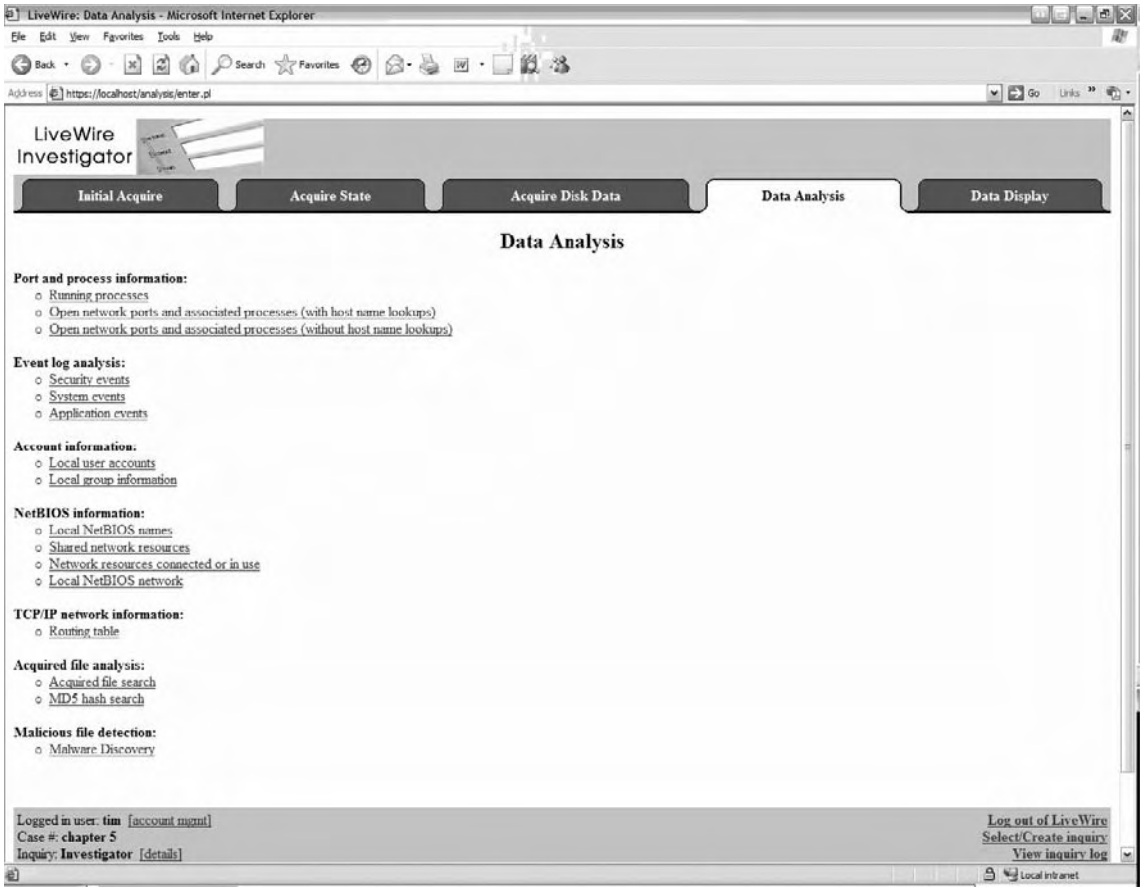
Live investigations allow investigators to capture volatile information that would not normally be present in a postmortem investigation. This information can consist of running processes, event logs, network information, registered drivers, and registered services. Why is this important to us, you ask? Let's take a look at the case of running services and how this could be extremely important us.

Running services tell us the types of services that may be running on a computer. These services run at a much higher priority than processes, and many users are unaware that these

services actually exist. Given their high priority and lack of attention by the typical end user, they are a common target for hackers. By conducting a live investigation, we are able to see the state of these services, which could prove crucial to our investigation. For example, a hacker could turn off the service for McShield, which is a McAfee Antivirus service, and then later come back and infest the machine with malicious software.

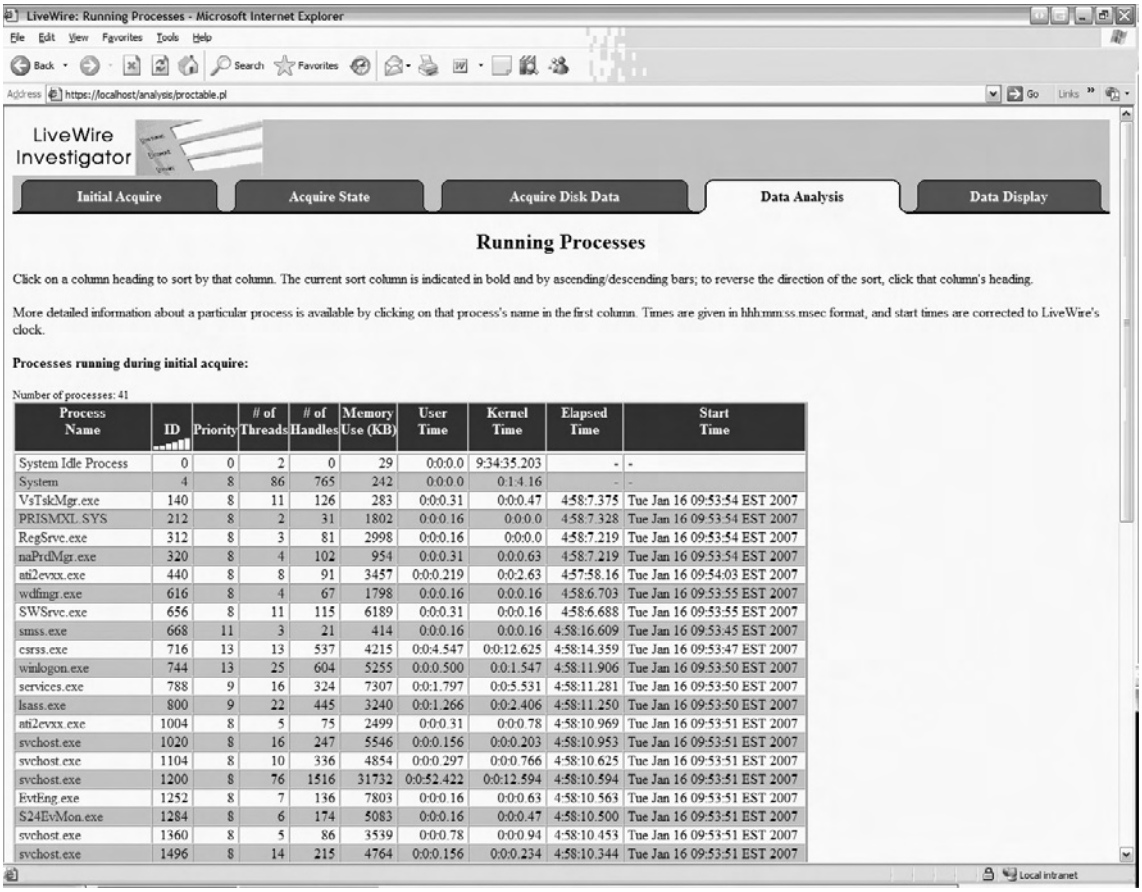
You might argue in the case of registered drivers that you could get a list of the drivers in a postmortem investigation. This is true; however, if you are at a crime scene and you conduct a live investigation, you might be able to see a driver for a digital camera. So you know to look for that camera in your surrounding area. But if you left the location, and then returned later to find that camera driver, you could only hope that the camera is still there when you make it back. As shown in the previous example, seeing registered drivers gives investigators knowledge of the peripherals of a suspect machine. Figure 7.11 illustrates some of the volatile information you can obtain about a systems state.

Figure 7.11 Shows an Example of Live System Information You Can Obtain Using Wetstone's LiveWire



Viewing running processes with the associated open network ports is one of the most important features of analyzing the system state. To peek into a system and correctly assess what processes are running and what ports they may be using is critical when trying to perform an investigative triage. Figure 7.12 offers a detailed look at the running processes of a target machine under investigation.

Figure 7.12 View of Running Processes Using Wetstone's LiveWire

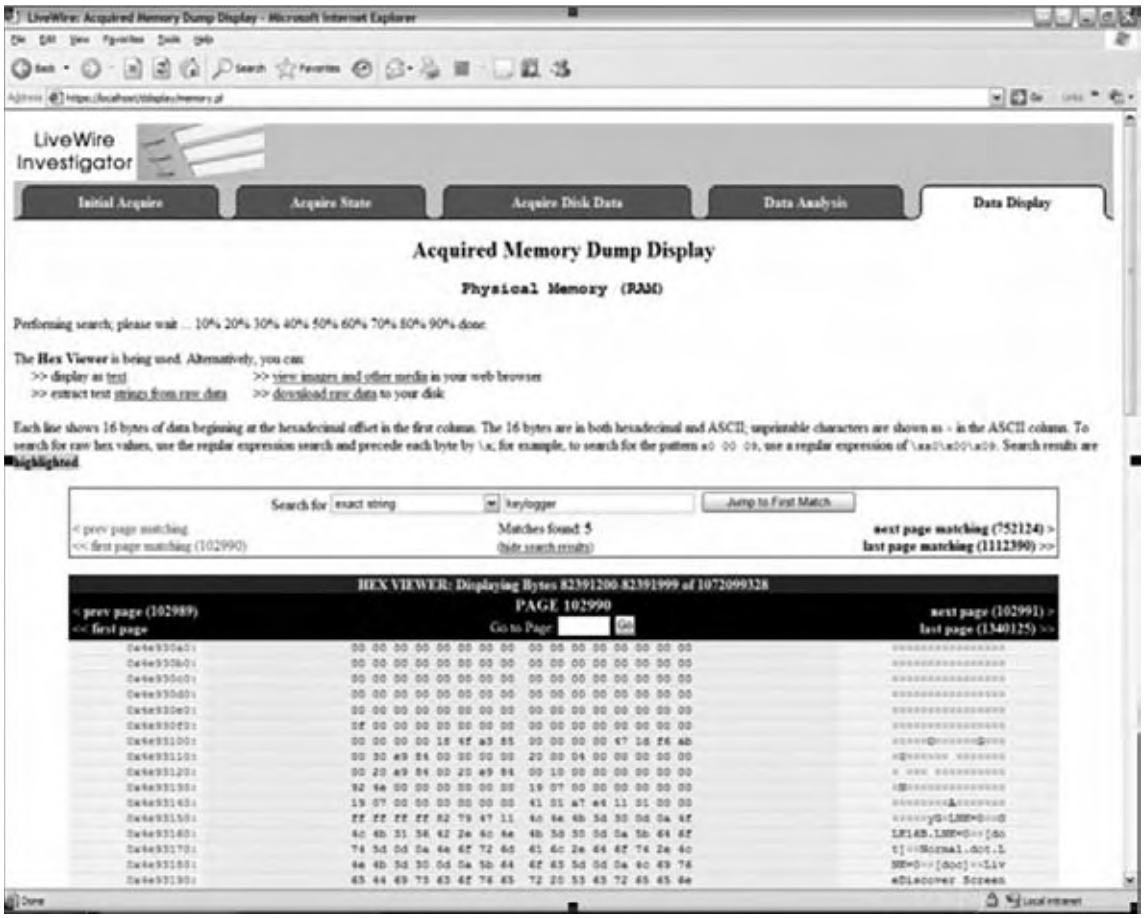

LiveWire: Running Processes - Microsoft Internet Explorer
File Edit View Favorites Tools Help
https://localhost/analysis/proctable.pl
LiveWire Investigator
Initial Acquire
Acquire State
Acquire Disk Data
Data Analysis
Data Display
Running Processes
Click on a column heading to sort by that column. The current sort column is indicated in bold and by ascending/descending bars; to reverse the direction of the sort, click that column's heading.
More detailed information about a particular process is available by clicking on that process's name in the first column. Times are given in hhh:mm:ss.msec format, and start times are corrected to LiveWire's clock.
Processes running during initial acquire:
Number of processes: 41

Process Name	ID	Priority	# of Threads	# of Handles	Memory Use (KB)	User Time	Kernel Time	Elapsed Time	Start Time
System Idle Process	0	0	2	0	29	0:0:0.0	9:34:35.203	-	-
System	4	8	86	765	242	0:0:0.0	0:1:4.16	-	-
VsTskMgr.exe	140	8	11	126	283	0:0:0.31	0:0:0.47	4:58:7.375	Tue Jan 16 09:53:54 EST 2007
PRISMXL.SYS	212	8	2	31	1802	0:0:0.16	0:0:0.0	4:58:7.328	Tue Jan 16 09:53:54 EST 2007
RegSrvc.exe	312	8	3	81	2998	0:0:0.16	0:0:0.0	4:58:7.219	Tue Jan 16 09:53:54 EST 2007
naPrdMgr.exe	320	8	4	102	954	0:0:0.31	0:0:0.63	4:58:7.219	Tue Jan 16 09:53:54 EST 2007
ati2evxx.exe	440	8	8	91	3457	0:0:0.219	0:0:2.63	4:57:58.16	Tue Jan 16 09:54:03 EST 2007
wdfmgr.exe	616	8	4	67	1798	0:0:0.16	0:0:0.16	4:58:6.703	Tue Jan 16 09:53:55 EST 2007
SWSrvc.exe	656	8	11	115	6189	0:0:0.31	0:0:0.16	4:58:6.688	Tue Jan 16 09:53:55 EST 2007
smss.exe	668	11	3	21	414	0:0:0.16	0:0:0.16	4:58:16.609	Tue Jan 16 09:53:45 EST 2007
csrss.exe	716	13	13	537	4215	0:0:4.547	0:0:12.625	4:58:14.359	Tue Jan 16 09:53:47 EST 2007
winlogon.exe	744	13	25	604	5255	0:0:0.500	0:0:1.547	4:58:11.906	Tue Jan 16 09:53:50 EST 2007
services.exe	788	9	16	324	7307	0:0:1.797	0:0:5.531	4:58:11.281	Tue Jan 16 09:53:50 EST 2007
lsass.exe	800	9	22	445	3240	0:0:1.266	0:0:2.406	4:58:11.250	Tue Jan 16 09:53:50 EST 2007
ati2evxx.exe	1004	8	5	75	2499	0:0:0.31	0:0:0.78	4:58:10.969	Tue Jan 16 09:53:51 EST 2007
svchost.exe	1020	8	16	247	5546	0:0:0.156	0:0:0.203	4:58:10.953	Tue Jan 16 09:53:51 EST 2007
svchost.exe	1104	8	10	336	4854	0:0:0.297	0:0:0.766	4:58:10.625	Tue Jan 16 09:53:51 EST 2007
svchost.exe	1200	8	76	1516	31732	0:0:52.422	0:0:12.594	4:58:10.594	Tue Jan 16 09:53:51 EST 2007
EvtEng.exe	1252	8	7	136	7803	0:0:0.16	0:0:0.63	4:58:10.563	Tue Jan 16 09:53:51 EST 2007
S24EvMon.exe	1284	8	6	174	5083	0:0:0.16	0:0:0.47	4:58:10.500	Tue Jan 16 09:53:51 EST 2007
svchost.exe	1360	8	5	86	3539	0:0:0.78	0:0:0.94	4:58:10.453	Tue Jan 16 09:53:51 EST 2007
svchost.exe	1496	8	14	215	4764	0:0:0.156	0:0:0.234	4:58:10.344	Tue Jan 16 09:53:51 EST 2007

Notice how we can see not only the process's name in Figure 7.12, but also the priority, the number of threads, number of handles, memory usage, and uptime. Again, you might ask why all of this is important. Well, if you are trying to assess what someone is currently doing, or even what they have done in the past, this information is critical. In addition, in the world of memory resident executables, analyzing the current process list is vital.

In a postmortem investigation, physical memory (RAM) is potentially the most important piece of evidence that is lost. However, this crucial piece of evidence is easily captured using live forensic and investigative tools, allowing the entire contents of RAM to be captured locally and even remotely. In Figure 7.13, we can see the contents of a memory dump and can conduct a search for the word *keylogger* in memory.

Figure 7.13 A Keyword Search for the Term Keylogger in a Memory Dump Using Wetstone's LiveWire



The raw data contents of the memory provide a vast amount of information that could have been lost if the machine was powered down for a postmortem investigation. Memory contains evidence ranging from user accounts, passwords, unsaved document content, and malicious software.

Terminology Alert

Malicious Software

Malicious software is a term describing a broad range of tools. However, memory-resident malicious software generally is seen with rootkits, Trojan horses, worms, and keyloggers. The following example contains a detailed explanation on how some memory-resident malicious software work.

Computer Analysis for the Hacker Defender Program

Hacker Defender is a popular rootkit that is capable of hiding processes, files, and even open ports. By default, when Hacker Defender is executed, it hides every file containing the prefix "hxdef." As a result, the file "hxdef100.ini," which is part of Hacker Defender, is hidden as soon as Hacker Defender executes. This file is then hidden from all users and even Windows Explorer itself. However, the file still exists in physical memory. Using live investigation techniques, you can take a memory snapshot and identify the file "hxdef100.ini" stored in RAM (see Figure 7.14). This same method can be used to reveal any file or process that Hacker Defender hides (see Figure 7.15). During a postmortem investigation, any files or processes hidden by Hacker Defender may not be accessible to the investigator. Figures 7.14 and 7.15 show evidence of the Hacker Defender program in the psychical memory of a computer.

Figure 7.14 Hacker Defender in Psychical Memory Using Wetstone's LiveWire

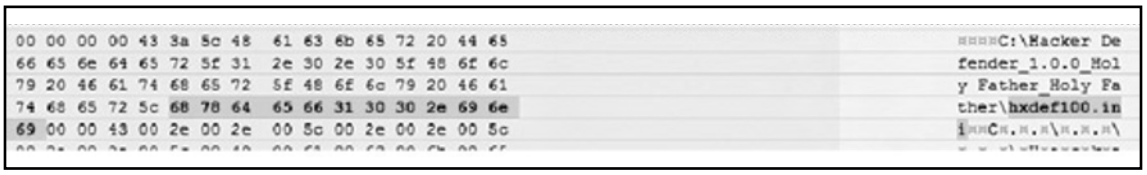


Figure 7.15 Another View of Hacker Defender in Psychical Memory Using Wetstone's LiveWire



As stated earlier, investigating a computer's system state is an important part of any investigation. It could help glean valuable information in a case and reduce the risk of missing data that could prove critical to your investigation.

Network Analysis

Often overlooked in live investigations is the environment in which the target computer resides. Data obtained from firewall laws, routers, intrusion detection systems, and so on are equally important to an examiner in obtaining the big picture. In the Hacker Defender case presented earlier, a defense attorney may argue that his client's machine was compromised and could not have committed the crime. A review of the firewall logs may show that the Hacker Defender activity from this computer was blocked, making this argument about the rootkit a moot point. As a live investigator, you should try to gain as much information about the network activity as possible. You might want to install a packet sniffer—with the appropriate permission, of course—and conduct a packet analysis of the traffic. Using this technique, you could determine if someone is connected to the box before conducting an analysis on the target machine. So remember, you may find additional evidence beyond the computer you are examining. Look for it.

Summary

As we move forward, computer forensics as we now know it will change dramatically. The release of Microsoft's Vista will include the ability for users to fully encrypt their hard drives. The use of virtual machines and virtual server farms are becoming more common place. Internet-based application servers will be harder for forensic examiners to physically collect. Additionally, Internet-based applications may generate diskless work stations, leaving the only evidence in physical memory. Finally, software vendors are starting to deploy a larger amount of software that securely deletes data because of identity-theft concerns. Because of these changes, and as I have pointed out in the preceding examples, I surmise that traditional forensics will become more impractical, and live investigations will become a necessity rather than a luxury Traditional methodologies are becoming somewhat obsolete. The need to adopt a new way of conducting these types of investigations is essential. While we have shied away from touching the computer in order to prevent any changes, it is now obvious that there are times when an examiner must interact with a live computer in order to retrieve vital data. Under the circumstances described earlier, you should be able to provide a reasonable explanation to any judge or jury as to why live forensics was used in place of traditional methods. However, should none of these circumstances exist, it may be best just to pull the plug.

Special Thanks

I would like to give thanks to my colleagues Christopher L. T. Brown and Chet Hosmer for their help with this chapter. There wisdom and insight into incident response and network forensic issues were invaluable.

References

Brown, Christopher L.T. *Computer Evidence Collection & Preservation*. 2006. Massachusetts: Charles River Media, Inc.

Chirillo, John. *Hack Attacks Revealed*. 2001. New York, John Wiley & Sons, Inc.

Mandia, Kevin et al. *Incident Response: Investigating Computer Crime*. 2001. California: Osborne/McGraw-Hill.

McClure, Stuart et al. *Hacking Exposed: Network Security Secrets & Solutions*. 2001. California: Osborne/McGraw-Hill.

Szor, Peter. *The Art of Computer Virus Research and Defense*. 2005. New Jersey: Addison-Wesley.

Frequently Asked Questions

Q: Can I view encrypted data in a live environment without having the password?

A: The answer is yes, provided that the drive or file is unencrypted on the suspect's machine.

Q: Can I view hidden processes like rootkits on a live computer?

A: Using special software, you can view hidden processes and files on a live computer.

Q: If I cannot image the entire drive, can I just copy the files I need?

A: Yes, you can copy the files you need using live forensic software to ensure you have the entire copy. Also, take notes when doing this since you may have to testify later about why you chose this method and what, if anything, you changed.

Chapter 8

Seizure of Digital Information

Solutions in this chapter:

- **Defining Digital Evidence**
- **Digital Evidence Seizure Methodology**
- **Factors Limiting the Wholesale Seizure of Hardware**
- **Other Options for Seizing Digital Evidence**
- **Common Threads within Digital Evidence Seizure**
- **Determining the Most Appropriate Seizure Method**

☑ **Summary**

☑ **Frequently Asked Questions**

Introduction

Computers and digital devices are employed by the majority of people in the U.S. for myriad business and personal uses. Because of the wide acceptance of computers in our daily lives, it is reasonable to conclude that people will use a computer to assist them in the commission of crimes, record aspects of crimes on a computer, and use computers to store the fruits of their crimes or contraband.

Any of the computers involved in the situations just discussed will likely contain upwards of hundreds of thousands of pieces of information stored in a digital format, including operating system files, program files, user documents, and file fragments in drive free space. While the challenge for the laboratory examiner is to find the relevant *data objects* on a hard drive or other media, a greater challenge exists for the on-scene responders and investigators: How can the information be collected from the scene and brought to a location where it can be examined? Does all the hardware on-scene need to be seized as evidence, or will an exact copy of the information serve the purposes of an investigation? Are there other seizure options to be considered?

Notes from the Underground…

Data Objects

Throughout this chapter, the term "data object" will be used frequently to discuss information found on a storage device or a piece of storage media (SWGDE, 2000). The digital information on a piece of media is nothing more than a long string of 1s and 0s recorded on either magnetic, solid-state, or optical media. Hard drives and floppy disks are examples of magnetic media; USB thumb drives and flash memory cards are examples of solid-state media; and CDs and DVDs are types of optical media. Any number of digital devices, including computers, cell phones, and iPods, will have operating systems and programs that arrange the 1s and 0s into a particular order to create images, documents, spreadsheets, music, and so on. For the purposes of our discussion, each of these discrete arrangements of information that are logically organized into something meaningful will be called a *data object*. The choice to use the term "data object" instead of the more frequently used term "file" is based on the fact that not all organized digital information comes in the form of a file. Information attached to a file such as a file header and metadata are not technically separate files, but can be culled out from the file as separate data objects. Other types of information

Continued

found on storage media are not files, but fragments of files left by the constant write and overwrite of information caused by the deletion of existing files and the creation of new files. For example, a certain amount of an old file may be left behind when a new file is overwritten in the same space—so-called file slack space. Still other types of informational fragments may include files and commands temporarily stored in the swap file or within the RAM itself. For these reasons, I believe it is more appropriate to call these organized pieces of information "data objects."

What we consider to be evidence has a dramatic effect on how we view the electronic crime scene. The current model of digital evidence seizure is focused on physical hardware, which is appropriate in most situations. However, as we move forward from this point in time, factors such as the size of media and full-disk encryption will impact the ability to seize all the hardware on-scene for later analysis at a forensics laboratory. Other options besides wholesale hardware seizure—RAM recovery, on-scene imaging of hard drives, and imaging of select files—need to become part of the basic toolkit of on-scene responders.

But the acceptance of other options for digital evidence seizure will not be a spontaneous event. The legal framework, the established workflows of existing computer forensic best practices, and the fear of the unknown will all play a part in determining how quickly the digital evidence seizure methodologies are adjusted to accept other options besides wholesale hardware seizure. The community of people that respond to, investigate, and prosecute crimes that have a digital evidence component is a very diverse population with different frames of reference and different technical understanding. If one group decides to unilaterally implement a change in practices or policy, the ripple effect is felt across the entire system—which is what makes *bridging the gaps* such an important part of considering and implementing any change resulting from advances in technology. As the author and a member of the greater crime-with-a-cyber-component-community, I hope this work serves to create discussion between the disparate communities on the appropriateness of both the familiar and innovative methods to seize digital evidence.

To these ends, I have organized the following pages to guide the reader through a number of topics relating to both the existing method of digital seizure and the innovative options available for on-scene responders. First, we will examine some of the framework surrounding the legal view of evidence, then we will address how the current digital evidence seizure methodology evolved, and afterward we'll take a look at each of the seizure steps individually. This work is not intended to be a step-by-step guide for digital evidence seizure, but many of the current best practices are examined, and some common pitfalls are discussed. Following the discussion of the current method of seizure, we will explore some of the reasons why the wholesale seizure of hardware on-scene may become problematic in

the future. Finally, we will discuss a number of options available for seizure of information, including the on-scene preview of information, the seizure of data held in the computer's RAM, on-scene imaging of entire hard drives, and the on-scene imaging of specific data objects.



WARNING

In the sections that follow, we will primarily be discussing criminal procedures, as I would hope that the civil procedures would follow the guidelines set forth by the criminal side of the house. Many civil procedures often turn into criminal events, and vice versa, so it's probably wise to be working each case as if it were destined for criminal court. Further, most of my work has been as a bridge between the technical community and that of law enforcement—and it is from this viewpoint that the chapter is written.

Obviously, criminals may actually steal a computer or other device directly—but the focus of this chapter is not on the physical theft of hardware. Instead, we target how information held within the storage medium can be processed into evidence.

Here, I will colloquially refer to computers and hard drives when discussing digital information. I do realize many types of digital devices and media contain data, but it is often too cumbersome to individually point out each item or specify each situation.

This chapter focuses more specifically on the seizure of digital evidence when that evidence relates to a static event, such as receiving a harassing e-mail or seizing a computer that contains child pornography. An analysis and discussion of recovering information and evidence from a more dynamic event, such as a Denial-of-Service attack or a network intrusion are covered in Chapter 7. Although much of what is discussed in the following sections still apply to network forensics, please note that I am purposely minimizing the points that apply to it.

Finally, I am not a lawyer, nor do I play one on TV. The intent of this chapter is to provide investigators, prosecutors and private sector personnel with options and discussion topics related to the collection of digital evidence. **Any conclusions or recommendations in this chapter that may resemble legal advice should be vetted through legal counsel. Always check with your local jurisdiction, local prosecutors, and local forensics laboratory as to their preferred method(s) of digital evidence collection.**

Defining Digital Evidence

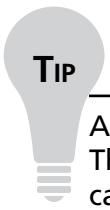
Black's Law Dictionary—the Bible for legal definitions—provides several definitions for *evidence* (Nolan, 1990). One of the definitions reads "Testimony, writings, or material objects offered in proof of an alleged fact or proposition." I have to say it is rather refreshing to have a generally straightforward and concise legal definition; generally, I don't equate straightforward and concise with legal…well… anything. The definition does provide a good launching point for our discussions on how digital information is viewed in the criminal justice system.

Black's definition of evidence as applied to digital evidence can be viewed in two ways. First, we can examine the computer itself as the evidence. This is clearly the case when the computer is the actual instrument of the crime, such as when the physical parts of the computer are used to commit a crime—for example, I hit you over the head with a keyboard. Colloquially, most law enforcement investigators and prosecutors will call the computer itself evidence even in cases where information on the computer relates to a given crime. As one investigator told me: "Everything seized at a crime scene is evidence until someone tells me it's not." In this sense, when the computer itself is seized at a crime scene or through a warrant, it is considered by many to be evidence.

Building on the view of the computer as evidence, many assert that the information on the computer requires the original computer to view the contents. In other words, the original computer—along the lines of how the best evidence rule requires the "original" whenever possible—may have an impact on how the information on the computer was actually viewed by the suspect. This is a valid viewpoint because many forensic software packages will not provide a view that is exactly as the suspect would have seen it. Too many different programs may show a given file, image, movie, or e-mail in a particular manner. The computer forensic analysis programs will often use a generic viewer capable of displaying any number of different formats. For example, Access Data's FTK has a generic format in which all e-mails would be displayed regardless of the program in which they were created. The generic format provides all the same information that would have been shown in the original e-mail, but it clearly is shown in a very different format than what the suspect would have seen. An e-mail viewed through the AOL e-mail program will include all the banners, advertisements, and formatting that make up the AOL look and feel of the user's experience. The e-mail itself will contain a number of standard fields, such as the e-mail header and the body of the message. The AOL program places these fields in a particular "package." However, that same e-mail viewed in FTK, though containing the same content, would lack the AOL packaging. In court, the examiner may be asked "Is this exactly what the suspect saw?" and the obvious answer is "No—but…" And it is within this "but…" that the court may suggest that the evidence—the complete computer and information as a unified package—be brought forth in front of the court.

A second way to view *Black's* definition is that the information, or *data objects*, contained on the digital storage medium are the "testimony, writings, or material objects" offered in proof of an alleged fact. This viewpoint makes the computer nothing more than a device that is used to access the information, and the components of the computer that store digital information nothing more than mere physical containers that house information—similar to a file cabinet or briefcase. Arguments can be made that only the desired information can be seized as evidence. The ramifications of this change in focus from hardware-as-evidence to information-as-evidence are far reaching.

If we do propose there is a distinction between the data objects and the physical container, we need to examine the legal framework within which we operate and seize information to determine if it is permissible to seize either the physical hardware or the information, or both. Rule 41 of the Federal Rules of Criminal Procedure (FRCP), titled "Search and Seizure" provides a definition for property, stating that "'Property' includes documents, books, papers, any other tangible objects, and information" (FRCP, Rule 41(a)(2)(A)). Within this definition is our first inclination that, in fact, the legal system views both storage containers and information as property. When we move forward in the FRCP into the discussions on seizure, we see that persons or property are subject to search or seizure and that a warrant may be issued for any of the following: (1) evidence of a crime; (2) contraband, fruits of crime, or other items illegally possessed; (3) property designed for use, intended for use, or used in committing a crime; or (4) a person to be arrested, or a person who is unlawfully restrained (FRCP, Rule 41[c]).



TIP

A number of legal documents will prove helpful in the coming discussions. The *Federal Rules of Evidence* (FRE) addresses the manner in which evidence can be presented in a federal court. *The Federal Rules of Criminal Procedure* (FRCP) provides the guidance for bringing an accused through the process of arrest and trial. The Computer Crime and Intellectual Property Section within the Criminal Division of the United States Department of Justice publishes a document titled *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations* (Manual). The Manual provides a very thorough review of a number of issues related to working with digital evidence—particularly as it relates to federal case law. Obviously, the depth of the information contained in the FRE, FRCP, and the Manual is well beyond the scope of this chapter, but I recommend that anyone interested in this field become familiar with these documents. Absent from the following discussions is talk of state law. Although many states will retain the ability for their own courts to be the "final say" regarding procedural or evidentiary matters, many states have adopted rules very similar to the FRE and FRCP.

Of interest to our discussion here is that property includes information, and that search and seizure is authorized, with a warrant, for property that is evidence of a crime. The next logical conclusion being that warrants can be issued for information that is evidence of a crime—but do the courts interpret using specific files or data objects as evidence, or should the focus be on the physical storage devices? Here, we consult the United States Department of Justice's Computer Crime and Intellectual Property Section's document titled *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations* (Manual):

> "The most important decision agents must make when describing the property in the warrant is whether the sizable property according to Rule 41 is the computer hardware itself, or merely the information that the hardware contains (pg. 61). …if the probable cause relates in whole or in part to information stored on the computer, the warrant should focus on the content of the relevant files rather than on the storage devices which may happen to contain them." The Manual references *United States v. Gawrysiak* (972 F. Supp. 853, 860 [D.N.J. 1997], aff'd, 178 F.3d 1281 [3d Cir. 1999]) which upheld the seizure of "…records [that] include information and/or data stored in the form of magnetic or electronic coding on computer media … which constitute evidence" of enumerated federal crimes (Manual, pg. 62). …The physical equipment merely stores the information that the agents have probable cause to seize. Although the agents may need to seize the equipment in order to obtain the files it contains and computer files do not exist separate from some storage medium, the better practice is to describe the information rather than the equipment in the warrant itself (pg. 65)…"

The guidance from the Manual is that the Rules on Criminal Procedure, and the interpretation of the same in the courts, points to the difference between the information held in data objects and the physical container (hard drive, flash media) in/on which the data resides. This provides some positive reinforcement to those that make the claim that the data itself is the evidence and that the computer or storage device is merely a vessel.

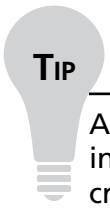
The preceding discussions regarding the computer as the evidence versus the data as the evidence has a dramatic effect on how we "seize" or "collect" evidence both at the scene and in the forensics laboratory. If your viewpoint is that the computer is the evidence, then your seizure methodology will be focused on the collection of the computer itself at the scene of the crime. If your viewpoint is that the information is the evidence, then you may be more inclined to attempt to locate and retrieve the information-as-evidence, with less care as to the eventual fate of the hardware. Further, you may be more inclined to call your "computer forensic" efforts simple "evidence collection" and remove the requirement for expert classification at trial. The important point here is that there are options to be considered, examined, and discussed within the community—options that have the ability to significantly change the entire approach to computer seizure and analysis.

Digital Evidence Seizure Methodology

The proliferation of personal computers changed how computers were involved in criminal issues. In the past, computers were often used primarily as the attack platform or target of the attack—now the more personal use of computes creates a situation where the computer is the storehouse of evidence relating to almost every type of crime imaginable. The result is that more computers are involved in some manner in crime and that more computers need to be examined for information of evidentiary value. But before they can be examined, they must be seized.

Previously, the highly trained computer specialist would attend to each seizure personally; however, the proliferation of computers and their use in criminal endeavors made personal attention to each case impractical. In some areas of the country, one specialist may serve an entire region. It is clearly unreasonable to believe that one specialist will be able to perform each seizure and complete the examination of the digital evidence for every crime with a cyber component. To fill this apparent gap in need versus capability, state and local law enforcement agents have become involved in recovering digital evidence from a crime scene where a computer is directly involved. Not only are state and local investigators faced with dealing with a new type of crime, but they are also asked to perform the seizures of digital evidence.

The on-scene responders/investigators often know very little about computers and often have not been instructed on how to "properly" seize digital information. Existing seizure protocols for physical items are used, resulting in a focus on the seizure of the computer hardware—sometimes the entire computer, including the monitor, printers, keyboard, and so on are seized and packaged for delivery to the lab. Over time, it became accepted to use the seizure methods focused on the seizure of the *physical hardware* for the seizure of *digital information*. Let's take a look at the flow of a general seizure of a personal computer.

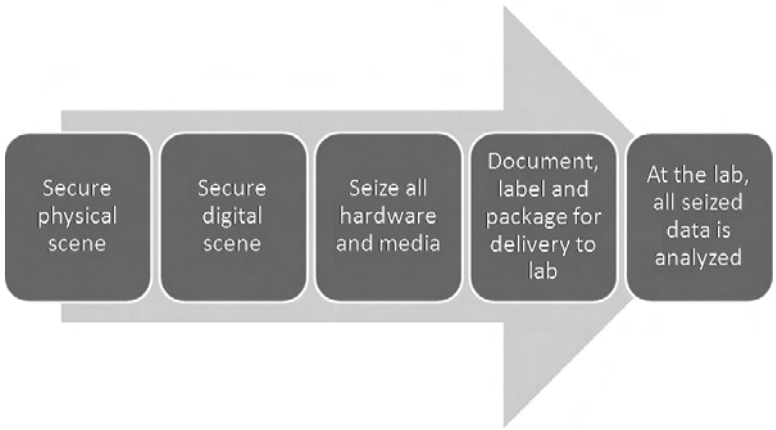


TIP

A number of other authors have nicely addressed the larger digital investigative model. Most notably, Carrier and Spafford present a "digital crime scene" model that exists within the physical crime scene (Carrier, 2003). Generally, these models present a complete framework for digital investigations, from incident response preparation right through to the examination and analysis of the seized information. Although this holistic viewpoint may be relevant to the administrator responsible for the entire operation, these models hold less applicability to the actual on-scene seizure of the relevant information, which is the focus of this chapter.

The current manner of seizure of computer hardware expects that the on-scene responder has a general knowledge about computers—to the level of "THIS is a keyboard, THIS is a mouse, THERE is no 'any' key," and so on. Better yet, the responder should have basic training on digital evidence collection, or, at the very minimum, be able to consult a guide on best practices, such as the *USSS Best Practices Guide* (USSS, 2006) or the *NIJ First Responder's Guide* (NIJ, 2001). Next, the responder would arrive at the scene, secure the scene physically, and begin to assess how the digital evidence is involved. The responder would take steps to secure the digital crime scene, which may include inspecting the devices for physical booby-traps and isolating the devices from any networks. The responder then seizes as many physical containers—physical media including hard drives, CDs, DVDs—as necessary to ensure the seized items reasonably include the information with probative value. The seizure of the hardware/physical containers involves labeling all wires connected to the computer or devices, and photographing the scene—paying specific attention to the labeled connectors. The physical items are seized, documented, packaged, and prepared for transport to an offsite facility for examination. At the offsite facility, possibly the local police agency or a state/regional forensic laboratory, the seized physical containers are examined for data objects with evidentiary value. If found, these data objects are usually included in a forensic findings report and are printed out or copied to other media and then provided to the investigator and prosecutors. Figure 8.1 outlines the steps of the traditional method for seizing computer hardware.

Figure 8.1 Traditional Seizure Methodology

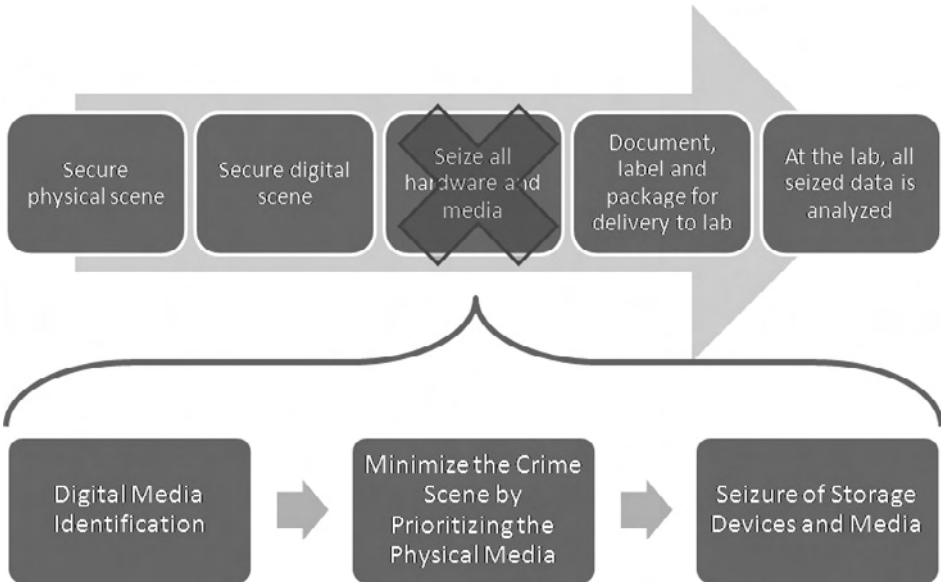

Secure physical scene
Secure digital scene
Seize all hardware and media
Document, label and package for delivery to lab
At the lab, all seized data is analyzed

That sounds pretty straightforward, doesn't it? For the most part, the preceding reflects the general process that the wide majority of law enforcement agencies follow when it comes to the seizure of digital evidence. As you can see, the general methodology reflects a focus on the seizure of the physical items. Further, the preceding model shows that a division exists between the investigators/on-scene responders and the forensic laboratory/examiners.

Seizure Methodology in Depth

Unfortunately, current seizure methodology does not adequately prepare our investigators to respond to scenes that are more complicated than a single machine sitting alone in a bare room. The fact is that the world is a messy place. Our responders need to understand that they need to have a methodology in place that allows them to work through more complicated scenes, such as finding dozens of computers or dozens of pieces of removable media or hundreds of CDs. *The steps presented in Figure 8.2 are representative of current seizure methodology, but the steps have been crafted to provide a higher level guidance about approaching non-standard seizure scenes.* Specifically, the "Seize All Hardware and Media" step shown in Figure 87.1 has been replaced by a series of three steps that help guide the responder through identifying all the digital media on-scene, minimizing the crime scene through prioritization, and then seizing the hardware and media that have the highest probability of containing the relevant evidence.

Figure 8.2 Seizure Methodology Featuring Minimization


Secure physical scene
Secure digital scene
Seize all hardware and media
Document, label and package for delivery to lab
At the lab, all seized data is analyzed
Digital Media Identification
Minimize the Crime Scene by Prioritizing the Physical Media
Seizure of Storage Devices and Media

We begin our seizure methodology at the scene, where a warrant for digital evidence is being served. It is assumed in the following that the scene has been physically secured, and the responder has a safe working environment. It is also assumed that the responder has a properly drafted warrant that identifies the information to be seized and outlines that an offsite examination of the media may be required if the situation makes the on-scene seizure infeasible.

Step 1: Digital Media Identification

The first step is to begin to canvas the scene in an attempt to locate the digital media that you believe has the highest probability of containing the evidentiary information described in the warrant. If the suspect has one computer sitting in his bedroom and another in a box in the attic, I'd bet my money that the information I'm after is the one in his bedroom. Taking a step beyond the simple situations, one needs to also consider removable media such as flash drives and CDs or DVDs. Flash drives are often held as personal file cabinets and may contain information of a personal nature. Look for flash drives on key chains, watches, in cameras, and just about anywhere—flash media can be unbelievably small. Another strategy is to look for media that contains backups of files from on-scene computer(s). If the information is important, you can be sure it will be backed up somewhere.

Where can digital media be found? The answer is pretty much anywhere. Locating very small, but very large storage media could be a significant issue when conducting a search. Be sure to balance the perceived technical expertise of the suspect versus the type of crime versus where you expect to find the relevant information. For example, it is fairly well documented that obsessive collectors of child pornography will gather tens-of-thousands of pictures of children being victimized. In this type of case, it would be most logical to be looking for a hard-drive or optical disks, given the amount of storage required. At this point in time, obtaining such large amounts of storage on flash media would be difficult, however. On the other hand, the same collector may be accused of taking pictures of children being victimized, and in this case the search should definitely focus on small flash media–type storage cards that could be used in a digital camera and/or be used to store and hide coveted images.

Documentation is part of every step, so this won't be the last time you see it mentioned. Nevertheless, it's worth mentioning here as a reminder. While conducting the search for digital media, it may be appropriate to narrate your movements into a voice recorder and to photograph the found media in place before moving it.

Step 2: Minimizing the Crime Scene by Prioritizing the Physical Media

After all the digital media is identified, an effort must be made to determine which storage devices or pieces of media have the highest probability of containing the information described in the warrant. Why? Because at some point it time, it will be impractical to seize all the digital devices, removable media, and storage media at a crime scene. At the current time, it may be possible to walk into a residence and only find one computer and maybe a few CDs. In this situation, the minimization of the physical media is all but done for you—you have in front of you only a few pieces of media that may contain the informational evidence. But technology is enabling homeowners to easily build rather complicated networks that may include wireless storage devices, multiple operating systems, shared

Internet connections, integration with traditional entertainment media, and integration with home appliances and devices. Downloadable and burnable movies and music are generally an accepted technology, greatly increasing the amount of optical media found in homes. Based on the availability of technology, on-scene responders will be faced with multiple computers, storage devices, and dozens to hundreds of pieces of media—all adding up to terabytes of information.

The responder must make some tough decisions about where she believes the information will most likely be found. One suggestion is to prepare a prioritized ranking to help decide which storage devices and pieces of media should be seized for offsite review. The prioritized ranking is also critical in deciding which devices or pieces of media are previewed on-scene—one of the options we'll be discussing later in this chapter.

Step 3: Seizure of Storage Devices and Media

The seizure itself is rather straightforward. After the scene is secured and it is determined that the hardware must be seized, the investigator begins by labeling all the connections/wires attached to the computer. Be meticulous in the labeling of wires and thorough in your documentation. It's a good practice to label both the end of a cable and place a matching label where the cable connects—for instance, label a Monitor's VGA Cable B^{1} and label the computer's VGA port as $B^{1'}$; label the monitor's power cable plug as B^{2} and label the wall outlet as $B^{2'}$. Photograph as many relevant objects and seizure steps as you see fit—digital photos are basically free and can be burned to disk and added to the case file. Don't forget to remove the sticky labels from the power outlets once they have been photographed.

After the computer has been labeled, documented, and photographed, disassemble the components and prepare the computer case for shipment. Best practices state that an unformatted floppy disk should be placed in the floppy drive with a piece of evidence tape sticking out like a flag. The presence of the disk in the floppy drive may prevent an accidental boot to the hard drive—but the new trend from computer and laptop manufacturers is to omit the standard floppy drives entirely, so this recommendation may be deprecated over time. Other options available to prevent an accidental boot are to unplug the power to the hard drive in a desktop machine and remove the battery from a laptop. Some recommend placing evidence tape over the external drives, including the floppy drive and any CD/DVD drives. When transporting, be careful not to drop, or otherwise jar or shock, the computer, as this may result in damage to the hard drive and possibly the motherboard. When transporting, keep the storage devices away from heat and strong magnetic fields, such as high-powered radios and big trunk-thumping subwoofers.



WARNING

Regardless of what hardware seizure methodology is written here or contained in any of the other published guides, always check with the laboratory or department that is going to process the seized hardware. Most have preferred methods for hardware seizure and transportation.

To Pull the Plug or Not to Pull the Plug, that Is the Question

I always wondered where the phrase *pull the plug* originated. I can picture a stressed out, overworked computer forensic technician on the phone with an on-scene responder, attempting to guide them through a proper shutdown and then a controlled boot process—prompting the following exchange:

Responder: It says to hit any key.

Forensic Tech: Uh-huh.

Responder: Hang on…. Um… where is the any key?

Forensic Tech: You've got to be kidding me…. Just pull the @#$@#% plug, wrap it in tape, and bring it to me!

Since that first hypothetical exchange—which still gives me a chuckle when I think about it—the mantra from the forensic community has been to pull the plug from the back of the machine, regardless of the state of the machine—on, off, writing to the drives, or anything else. *I have no doubt that, across the board, the simplest most teachable method of seizure that will generally preserve most of the data and evidence is to pull the plug from the back of the machine*. Pulling the plug and prepping it for transfer to an examination lab is the only option that is reasonably teachable in a few hours to first responders of any skill level. But, surely, we need to be able to do something other than pull the plug. We cannot possibly make advances in this field if we limit all officers and agents to a methodology based on the lowest common denominator.

The most pressing issue relating to *pull-the-plug* is that some operating systems (OSs) really like to be shut down properly. Rapid power loss in some OSs can actually corrupt the operating system's kernel or the central module of the system. Unix, Linux, and Macintosh operating systems are the most vulnerable, but some Windows-based OSs, such as a Windows 2000 server, should be shut down properly. Moore (2005) presents a good review of the proper shutdown method (shutdown versus pull-the-plug) for different operating systems based on the operating system's ability to recover from rapid power loss.

Obviously, if you intend to shut down the machine properly, you must determine the OS. To determine the OS and to initiate a proper shut down sequence, you need to manipulate the computer's mouse and/or keyboard, but manipulating the mouse/keyboard will change data on the suspect's machine. You say "But I'm not allowed to change data on the suspect's machine!" That may be the guidance given, but it is more appropriate to take the position: "I will do the most appropriate and reasonable actions during seizure to ensure I retain as much of the relevant information as possible. Here is the documentation of my actions." The focus here is on reasonableness and the documentation of actions. Also, it is important to key-in on the retention of the *relevant information,* which includes the information of potential evidentiary value and should not include the Registry changes made to indicate that a shutdown occurred. Simply put, moving the mouse to determine the OS and starting a shutdown sequence did not place 5,000 images of child pornography on the computer's hard drive. However, pulling the plug on a Linux system may actually impact the ability to recover those same images.

There is no one correct answer to the pull-the-plug question. If you have the skill and knowledge to determine the operating system of the suspect computer and you determine that the operating system and other data could be damaged by pulling the plug, then shut the machine down properly. Document your actions and explain clearly and knowledgeably how you prevented damage to the computer, and possibly to the evidentiary information, by following a shutdown procedure. Show how your actions preserved the evidence, as opposed to corrupting it. If you have the skill and document the steps you followed, you have solid footing on which to defend your actions. If you do not possess such skill, or if the more advanced techniques are not working in a given situation or on a particular piece of hardware, then by all means, pull the plug.

Factors Limiting the Wholesale Seizure of Hardware

Earlier we contrasted the historic seizure context versus the current context and discussed how the historic context placed a focus on the on-scene seizure of data objects, as compared to the current situation where the focus of the on-scene activities is to seize all the physical containers. The question I pose to you is this: Are we heading in the right direction by focusing on the seizure of the physical hardware (the container items) rather than focusing on the seizure of the relevant information (data objects)?

Earlier seizures of digital evidence focused on data objects because it was impractical to attempt to image an entire server, based on the high costs of storage media. I suggest we are heading toward a similar impracticality—although this time our inability to seize *all* the information is based on a number of different factors, including massively large storage arrays, whole disk encryption, the abundance of non-evidentiary information on media and

related privacy concerns, and the time involved in laboratory forensic analysis. At some point in the future, the process by which we image entire pieces of media for forensic analysis will become obsolete (Hosmer, 2006).

I suggest we make the distinction that there other options beyond wholesale seizure available to our responders. We need to train our responders to have the ability to perform on-scene data preview, full data-image, and imaging of only the relevant data objects. Further, we need to begin to change the wholesale seizure paradigm now—for all responders not just the specialists—before we are faced with a greater volume of cases we are ill prepared to address.

Factors Limiting Wholesale Seizure: Size of Media

Storage devices are getting big—very big. Now, at the end of 2006, it is quite common for a single hard drive to contain 100 gigabytes of information—roughly equivalent to a library floor of academic journals. It is very achievable for the home user, both technologically and financially, to put together a 2-terabyte storage array—an array that could house the complete works within an entire academic research library (SIMS, 2003). Storage is relatively cheap, and people are taking advantage of the extra space by storing music, movies, and creating mirrored backups (RAID 1 arrays). The typical crime that involves a computer won't include a multi-hundred terabyte server, but showing up at a crime scene with a 200-gigabyte destination drive and finding a 1.5-terabyte RAID will certainly have a negative impact on your ability to create an on-scene image of the data.

What exactly happens when the full 1.5 TB RAID and 200 DVDs are seized and brought back to the forensic laboratory for analysis. Do you actually have the hardware and software to acquire and process that much data? If the laboratory is not a regional or state lab, but a small laboratory set up at the local agency, the answer might be yes—but processing the case might use the entire budget set aside for target drives for the entire year for that one case. Once the data is examined, does the jurisdiction or local policy dictate that the imaged data be archived? At some point, the ability to seize and process *everything* will exceed the budget set aside for the purchase of forensic processing computers, target drives, and archival media and will also exceed the time available for forensic examiners to process the case.

Factors Limiting Wholesale Seizure: Disk Encryption

A number of encryption programs exist now that provide whole disk encryption, a common one being PGP from pgp.com. These types of encryption programs encrypt all the data on the hard drive and are generally transparent to the user; meaning that one password in the startup sequence "unlocks" the contents for viewing and editing. Of course, looming on the

horizon is the Windows Vista operating system, purported to incorporate BitLocker Drive Encryption tied to the Trusted Platform Module cryptographic chip in the higher-end versions of the operating system.

Whole disk encryption has some serious implications for law enforcement when performing seizures. First, if a whole disk encryption is enabled on a running computer, and the computer is shut down or the power is removed, there is a very good chance that the data on the drives will be unrecoverable without the proper key. Responders may need to determine if a whole disk encryption program is enabled before shutting down / pulling-the-plug on a computer during seizure. If one is present, bringing the computer back to the lab for analysis may be futile. One of the best chances to retrieve the evidentiary information is when the machine is running and the user has access to the files. Second, the implementation of the TPM chip may lock the drive so the data may only become available on a specific machine. This would prevent an image of the drive from being booted in another computer or viewed with a computer forensics program. The use of disk encryption is forcing law enforcement to have other data seizure options available beyond the seizure of physical hardware.

Factors Limiting Wholesale Seizure: Privacy Concerns

Personal computers often contain myriad information about a person's life, including financial, medical, and other personal information, information related to their job (such as work products), and even information owned by several people, possibly a spouse, family member, or roommate. It's unclear how the criminal and civil courts would view a challenge from an impacted third party regarding the seizure of a common computer. However, if that third party maintained a blog or Web site, their information may be protected from seizure under the Privacy Protection Act (PPA) (42 U.S.C. § 2000aa). The PPA was specifically developed to provide journalists with protection from warrants issued to obtain information about sources or people addressed in their publications. The PPA reads "…it shall be unlawful for a government officer or employee, in connection with the investigation or prosecution of a criminal offense, to search for or seize any work product materials possessed by a person reasonably believed to have a purpose to disseminate to the public a newspaper, book, broadcast, or other similar form of public communication." The PPA may not protect the person that possesses the information if that person is suspected of committing the criminal offenses to which the materials are related. Simply put, if you committed a crime and you have publishable information related to that crime on your computer, that information most likely will not be protected under the PPA. However, the PPA may protect the interests of a third party that uses or stores data on a computer, and may possibly protect the information of the accused if the information does not relate to the crime being investigated.

The potential situations of co-mingled evidentiary data and publishable materials, each owned by a separate person do sound unlikely if you only consider a single computer. But what if you consider a network addressable storage device located in a home network? For example, let's say that such a storage device exists at the scene of a seizure. Every member of the household stores information on the device, and little Susie's unposted blog entries on her life-as-a-brainy-15-year-old-girl are located on the storage device commingled with the information described in the warrant. Although you may seize the storage device, you may also be involved with other court proceedings related to the violation of the PPA—civil, and possibly criminal, proceedings where you are the defendant!

The Secret Service ran across a similar situation in the case of *Steve Jackson Games, Inc. v. Secret Service* (*Steve Jackson Games, Inc. v. Secret Service*, 816 F. Supp. 432 [W.D. Tex. 1993]). The Secret Service seized two computers from the company, believing that the company's system administrator had stored evidence of a crime on company computers. The day after seizure, the Secret Service learned that the computers contained materials intended for publication; materials that belonged to the company. Regardless, the Secret Service did not return the computers until several months had passed. The district court ruled that the Secret Service had in fact violated the PPA and awarded Steve Jackson Games $50,000 in damages and $250,000 in attorney's fees. The story of this raid goes well beyond the short summary provided here. The raid and the trial play a significant role in hacker mythology and also played a part in the formation of the Electronic Frontier Foundation (Sterling, 1994). Nonetheless, the moral of the story is that the Secret Service was not prepared to seize the specific information described in the warrant when they learned of the to-be-published materials present on the seized hardware. It's not known how the Secret Service would have changed their seizure methodology if they knew about the publishable materials before they served the warrant—but, for example, if they didn't have the capability of solely seizing the relevant data objects, the Secret Service might have had no other option but to seize the hardware. This example goes to show that having other seizure options available may be a critical skill that determines the success of an investigation.

Factors Limiting Wholesale Seizure: Delays Related to Laboratory Analysis

If investigators of crimes involving a computer rely completely and absolutely on their computer forensic laboratory for the processing of their seized hardware in search of evidence, they are at the mercy of the timing dictated by the laboratory. From my experience, a computer forensic laboratory can process anywhere from 30 to 60 cases per examiner per year; possibly more depending on the types of cases they work and their equipment, but considering most forensic laboratories are government agencies, I doubt they are operating year after year on the most current computers available. To make matters worse, the increase

in the size of storage media has far outpaced the increases in processor power. The same $500 that could afford a 100 MB drive in 1991 can now put a 750 GB drive in your pocket. Compare that to a 50-MHz Intel from 1991, next to a 3-GHz processor in today's fastest computers, and you'll see that the cost effectiveness of hard drives grew 125 times faster than that of processors from 1991 to the present (Gilder, 2006). Depending on the backlog at the laboratory, investigators can be faced with waiting up to—and over—a year for the results of their examination to be returned from the lab.

I am unable to specifically quantify how delays in the forensic examination are impacting investigations and prosecutions, but I can offer my opinion that delays in the processing of digital evidence are one of the most significant impediments in investigations and prosecutions that have a digital-evidence nexus. Given the opportunity to perform an on-scene seizure of the relevant information versus being forced to wait one year for the results from the laboratory, the choice will be clear for many investigators. However, there are difficulties and challenges in seizing the information on-scene—but these challenges must be weighed against the time delay in receiving the processed evidence.

One investigator I interviewed about this type of situation described a child pornography possession case where there was a chance that the accused possessor was also creating and distributing images of child sexual abuse. Unfortunately, the investigator had no means to preview the digital information on-scene, nor back at the department, nor did the investigator have the ability to perform a digital information analysis in-house. The computer was sent off to a computer forensics laboratory, where it sat in the queue behind other just-as-important cases. Because the information could not be reviewed, the investigator had no evidence to substantiate the drafting of an arrest warrant for either the possession of child pornography or the child sexual abuse. In such cases, any delay caused by a backlog at a forensics laboratory not only impacts an investigation, but also has a direct effect on a (potential) victim and continued victimization.

Protecting the Time of the Most Highly Trained Personnel

Digital devices have become almost completely ubiquitous in our current society. The legends of "convergence" are slowly coming true, where the line between computers, cell pones, cameras, and so on is now fuzzy and may disappear altogether in the future. IPv6 looms on the horizon and promises to equip every device, from cars to toasters, with an IP address. How do we find the time to train our law enforcement community in an entirely new set of skills? What is the balance between knowing enough and making a specialist out of everyone?

Determining whether the individual data objects with evidentiary value are seized or the storage media is seized will likely depend on the technical prowess of the responding investigator. The best situation would be to have a team of highly trained digital evidence

seizure specialists respond and then properly prepare a Windows computer for seizure. The reality is that there will never be enough computer specialists to respond to every crime scene—let alone a "team" of them—to seize every piece of information or computer involved either directly or peripherally in a crime.

Looking forward, we can anticipate that the number of computers and other electronic devices requiring seizure and examination to surely increase. Clearly, from all accounts of the situation, the current methodology has its flaws. Delays in the examination of seized digital media are frustrating investigators and are impacting prosecutions. Although we clearly need more computer forensic specialists, do we have the resources—specifically the personnel, time, and money—to train and equip enough specialists to meet the current demand for seizures and exams? What about future demands? From what I have observed, I don't believe we have anywhere near the number of qualified personnel to address the current issues, let alone what the future will hold. Nor do I believe that the existing infrastructure can support the required increase in the number of computer forensic examiners or specialists. Most agencies fight for the addition of a single position—so I'm doubtful that the system will suddenly change and begin hiring scores of new personnel.

The situation comes down to a simple law of economics: productivity will only be increased by adding more people or making existing people more efficient. We don't really have the ability to throw more people at the problem, so the only option is to do more with the people we have. As it pertains to cyber crimes and crimes with a high-technology component, this means we cannot continue to rely on computer specialists for every aspect of an investigation that involves a computer. Every law enforcement agent, from on-scene responders to detectives performing investigations, now have a duty to begin to pick up the slack that has created the conflict between the large—and growing—number of crimes with a high-technology component and the relatively small number of specialists available to work these types of cases. We need to consider the computer specialists and the computer forensic laboratories as a finite resource, and any constructive work performed in the field by patrol officers or detectives reduces the strain on the forensic system. *With this view, the most valued resource is the time of the highest-trained individuals*.

The general scenario of protecting the time of the most highly trained individuals so that they may focus on the most important issues is not a new concept. Those trained in hazardous material response work under a pyramid-like distribution of knowledge; the wide base of the pyramid consists of awareness-level trained people, while the small tip of the pyramid consists of highly trained specialists. Not only are these training levels generally accepted within the hazardous material response community, but they are codified in 29 CFR 1910.120(q)(6). The training code establishes the general level of knowledge, the hours of required training, and what can be expected from responders that have achieved each of the training levels. Because the different training levels are clearly defined, each responder on-scene understands their role and, more importantly, the role of other responders. Those with awareness-level training are taught to basically recognize that something bad has happened, call for help and watch from a

distance with binoculars. Operations-level training prepares responders to respond in a defensive fashion, without attempting to stop the release. Technician-level responders are trained to attempt to stop hazardous material release, and specialist-level responders usually have specific knowledge pertaining to a particular chemical. At each level, the responder receives more training to be better prepared when responding to a scene.

At the current time, it would not be practical to attempt to regulate or codify the training requirements or duties of those involved in digital evidence seizure, but it is important to recognize that people of different training levels will likely approach seizure in different ways. (See Figure 8.3.)

The seizure methodology that is developed for the knowledge level of the non-technical responder is in direct conflict with the best possible seizure scenario. Any seizure methodology adopted by an agency must be fluid enough to allow a minimally trained responder and a highly trained responder to both seize the digital information in the manner most applicable to their knowledge level.

Figure 8.3 Digital Evidence Seizure



The Concept of the First Responder

Who exactly is the "First Responder" referenced in numerous digital evidence seizure guidelines and reports? Is the first responder simply the person that happens to be on-scene first? If yes, then the first responder could be any line officer. If every first responder needs

to be trained to seize digital evidence, and we acknowledge that the seizure methodology will be necessarily fluid based on the responder's technical knowledge, you begin to see the problems involved with designing one particular training for first responders.

A second issue is the number of hours of training that could be allotted for first responder training. Will the administration of an organization allow their personnel to take a half-day course on digital evidence seizure? Probably. Realistically, though, what could you cover in four hours of instruction? I would guess the limit would be the recognition of digital evidence. So, would a two- or three-day training be sufficient to cover the recognition of digital evidence plus the seizure of digital information? Possibly, but would the people attending that training still be considered first responders or would the additional training necessitate they become specialists in this area? I am doubtful an agency's administration would agree to send every line officer to a three-day training to be first responders.

We are clearly caught in a catch-22. All line officers need to be able to seize digital evidence, but the first responder–level of training may not fully equip the officers to seize the evidence. The level of training required to more completely understand the digital evidence seizure process may involve multiple days of training, and multiple days of training on a single topic will most likely not be provided to all line officers. Unfortunately, it is not as simple as identifying one cadet in the academy that will specialize in investigating crimes with a cyber component, and putting this cadet through weeks of specialized training. The ubiquity of computers and digital evidence make the training of one single person insignificant—everyone's expertise needs to be raised to allow the specialists to focus on more technically challenging crimes.

There will be no clear-cut answer to this dilemma, but a number of factors could help mitigate the issue. First, law enforcement officers need more training in general computer skills. During a law enforcement officer's daily work, which is more likely? Arrest a suspect, be involved in a shooting, or spend some time working at a computer? The answer is a no-brainer—computers are an integral part of the law enforcement landscape and most officers cannot go a day without having some level of mission-critical interaction with a computer. However, the general level of computer knowledge among law enforcement personnel is low, and use of a computer is rarely a focus of academy setting. Providing law enforcement with basic, fundamental computer skills would not only impact their views toward digital evidence, but would also positively impact their daily work activities.

Second, all law enforcement personnel should receive basic awareness–level training on digital evidence. Awareness-level training need only cover the basics of a computer and where digital evidence may be stored. It is important for all officers to recognize that storage media, particularly flash-based media, may be no larger than a postage stamp, yet possibly contain several gigabytes of information. Understanding that many seemingly single-purpose devices, such as cell phones or mp3 players, may contain other types of information—for

example, documents may be stored on an mp3 player—will have important investigative implications far beyond simple search and seizure concerns. Perhaps the next time a drug dealer is arrested with a PSP, you may want to search him for a small flash media card—as a dealer, his contact list might be accessed from the flash card on the PSP. Until a more uniform level of basic knowledge and awareness is reached among law enforcement, it is hard to speculate how the increased awareness will benefit investigations. But as the saying goes, you miss 100 percent of the shots you don't take, and more appropriately, you miss 100 percent of the evidence you don't look for.

Third, any seizure methodology developed and/or adopted by an agency must be fluid to allow for seizures to be conducted by both minimally trained individuals as well as highly trained specialists. Do you want to put your specialist on the spot when he breaks protocol to perform a function that is technically more appropriate? Conversely, do you want the specialist to be on-scene at every warrant service, arrest, or vehicle search? There must be options within the methodology that allow each officer to act reasonably according to their skill level.

Other Options for Seizing Digital Evidence

The wholesale seizure of the physical storage device/media is arguably the most common form of seizure practiced by law enforcement responders today. The question remains, are there other options besides the seizure of physical devices that are available to responders? If yes, are these methods of seizure within the reach of anyone but the most technical of responders?

For a long time, up to and including today, many in the forensics community place little faith in the ability for responders on-scene to deal appropriately with the computers they may encounter. The direction was simply "Don't touch the keyboard. Pull the plug and send everything to the lab." In many cases, the forensics side of the house is correct to protect against the possible corruption or destruction of data by taking this hard-line approach—particularly based on the technology of yesterday—but at what cost? Although the computer forensics community may have intended to do the most good by promulgating the pull-the-plug mantra, we need to examine how dis-empowering the on-scene responders may affect the overall forensic process, from seizure through analysis to investigation and ultimately prosecution.

The latest *Search and Seizure of Computers and Obtaining Digital Evidence* (Manual), published by the Department of Justice supports the proposition that the seizure of digital evidence should be an incremental process, based both on the situation and the training level of the responder. The Manual describes an incremental approach as a search strategy (pg. 221) for the seizure of digital evidence from a functioning company where the wholesale seizure of all the computers from the company would be impractical.

The Manual provides the following steps in its incremental approach:

1. After arriving on-scene, Agents will attempt to identify a systems administrator or similar person who would be willing to assist law enforcement in identifying, copying and/or printing out copies of the relevant files or data objects defined in the warrant.
2. If there are no company employees available to assist the Agent, the Agent will ask a computer expert to attempt to locate the computer files described in the warrant and will attempt to make electronic copies of those files. It is assumed that if the Agent is an expert, he/she would be able to proceed with the retrieval of the evidence.
3. If the Agent or expert are unable to retrieve the files, or if the onsite search proves infeasible for technical reasons, then the next option is to create an image of those parts of the computer that are likely to store the information described in the warrant.
4. If imaging proves impractical or impossible for technical reasons, then the Agent is to seize those components and storage media that the Agent reasonably believes includes the information described in the warrant.

The Manual has a focus on Federal law enforcement and the incremental search strategy is described in the context of responding to a functioning business where evidence of a crime may reside on the business's systems—hence, the focus in the Manual on gaining assistance from the business's systems administrator. Even though, realistically, you are not going to ask the suspect for help in retrieving the files of interest, there is good reason to expand this incremental search strategy to the search and seizure of digital information that resides on non-business systems. First, many home users set up networks similar to what would be present in a small business. Second, the amount of storage on a home network may exceed the amount of storage used for business purposes, as home users are more likely to possess large music and movie files. Lastly, current and impending technologies such as whole disk encryption make the offsite analysis of storage media impractical, if not impossible. A mechanism must be developed now that enables responders to pull evidence off of a running system before these types of systems are in widespread use. Otherwise, we may be changing the paradigm a few years too late.

Although the change in focus from hardware-as-evidence to information-as-evidence may be a radical departure from how many people currently view digital evidence, it is not exactly a new viewpoint. In fact, the change to a focus on the information as evidence may be a renaissance of sorts; the computer crime investigators of yesterday knew nothing other than the retrieval of relevant information from servers and networks. Much of the investigation of computer crime in a historic context related to examining events that occurred

within a network infrastructure. In his book from the pre-World-Wide-Web year of 1990, *Spectacular Computer Crimes*, Buck Bloombecker discusses numerous computer crimes, most of which involve attacks on the network infrastructure (virus, worm) or schemes that were enabled by the presence of a network infrastructure, such as stealing unauthorized computer time or manipulating the wire transfer system to steal bank funds.

Crimes with a cyber component changed dramatically following the personal computing revolution, which was hand-in-hand with the rise of the World Wide Web. Prior to the 1990s, few people with personal computers used them solely for personal purposes. Prior to the 2000s, few people were providing personal information about themselves for the world to view. So it's not surprising that when we take a look backward, we see that the investigation of cyber crime involved incident response tasks, like pulling logs and records off of servers and other infrastructure-level digital devices, and less often concerned the seizure of a personal computer. Wholesale duplication of servers was impractical, storage costs were high, and so it was cost prohibitive to attempt to pull together the necessary equipment to image the entire server. Although the investigators of the time were breaking new ground, they knew enough to document their actions, make best efforts not to change the data objects with evidentiary value, and image the relevant data objects so they could be printed or referred to at a later date. Responders to network intrusion events were faced with no other option but to seize the relevant data objects—which is still the case today.

Responding to a Victim of a Crime Where Digital Evidence Is Involved

There is an old saying that all politics are local politics. Although I'm not quite convinced of the particular weight of that adage, I do believe that all crime is local crime. The Internet may have created a global community, but crime, even crimes committed over the Internet, will be reported to a local agency. It is imperative that local agencies have the ability to field a complaint regarding a crime with a cyber component and be able to respond appropriately. I have heard horror stories where complaints of e-mail harassment, auction fraud, and other crimes with a cyber component were just ignored by a local agency. Yes, a statement was taken and a report prepared, but no follow-up investigation was conducted. Worse, I have heard of agencies telling victims that the investigation of their complaint involved the seizure of their machine for forensic analysis, and that the analysis might take over a year to complete. I think it's pretty obvious why the complaint was dropped.

The unfortunate part of the situation is that the responding officer (or local agency) places an improper focus on the technology and loses sight of the crime that occurred. Often, the technology used is secondary and of little relevance. It could be quite possible that harassing statements in an e-mail might be coming from someone the victim already knew. If the harassment occurred through some other non-seizeable, non-virtual means (for example, spray paint on a car), the officer would most likely follow up with a

knock-and-talk with the suspect. The follow-up on the e-mail harassment should use the same logic. Does the investigation need to be focused on tracing an e-mail to its source when you already have a good idea as to who sent the e-mail? It is important that investigators do not switch off their investigative skills because a computer is involved.

When responding to a victim, the focus must be on having the victim provide the law enforcement officer with something that substantiates their complaint—a print-out of the harassing e-mail with full header information, a cut-and-paste printout of the IM conversation where their child was sexually solicited, or a screen-print of a disturbing Web page. Any information that can be provided by the victim to a responding officer will increase efficiencies in the entire investigative process. The officer will be able to read the e-mail header and get preservation orders out to the ISPs; the detectives will be able to begin working the case, rather than securing another statement from the victim; and the computer forensics system won't be burdened by yet another machine requiring examination—particularly for data objects that could have reasonably been obtained on-scene.

Cases occur where the victim's computer must be seized. Harassments in e-mail or chat (when logging) that violate a protective order may have to be seized, depending on the situation. If a spouse or roommate finds child pornography on a computer, the computer should be seized since it contains contraband. But barring these unavoidable circumstances, the seizure of victim computers is often unnecessary and contributes to the logjam at the digital forensic laboratories.

When communicating with a victim, be sure you let them know to not delete anything on their system until their complaint has gone through the entire process. Also be quite sure to document the steps the victim took to provide you with the substantiating evidence. If you had to assist the victim in any way—maybe you showed them how to see full headers on an e-mail, for example—make sure those actions appear in the documentation. Make a note of the system time on the computer, and verify that the evidence contains a time and date stamp, and that the time and date make sense to the victim. Lastly, be responsive to the victim's needs. Many crimes with a cyber component—particularly frauds and thefts—will have an international component that makes the apprehension of a suspect and reimbursement to the victim nearly impossible. Be sympathetic and provide the victim with any resources that can assist them in dealing with banks, credit card companies, and creditors, such as a properly written police report. They have already been victimized; don't let your actions lead to a prolonging of the victimization.

Seizure Example

Here we will examine an example of a digital seizure to help explore the options available to on-scene responders. Let's start by saying that Sally receives a harassing e-mail from an anonymous sender. She believes it is a former co-worker named Sam, who has harassed Sally using non-computer-based methods before. The officer follows the guidance discussed in the

"Responding to a Victim of a Crime Where Digital Evidence Is Involved" section and instructs Sally to print off a copy of the e-mail showing the full header information. Sally prints off the e-mail as substantiating proof to back up her complaint, and the officer leaves the scene with a statement from Sally and a copy of the harassing e-mail.

You notice that Sally was not told that her computer would need to be seized and held for a year—which would, in effect, cause Sally to drop her criminal complaint and also drop her opinion of the police. Instead, the officer leaves the victim scene with a statement, and some level of proof to back up the complaint, which allows the investigation to proceed without undue hardship to the victim.

The investigator then uses the information contained in the e-mail header to contact the e-mail provider, legal paperwork is sent to the provider looking for the account holder's information, and finally the e-mail is traced back to Sam's Internet service provider (ISP) account. We now have a general confirmation that the e-mail was sent from a computer connected to Sam's ISP account—although this could be any number of computers at Sam's house and possibly even be a neighbor using Sam's wireless access.

The investigator drafts a search warrant affidavit looking specifically for the information that is relevant to this case—specifically a preserved copy of the sent e-mail. The investigator is careful to focus the search warrant on the information to be seized, and does not focus on the containers or storage media in which the information may reside. The investigator further notes that an incremental approach will be used, which dictates that onsite seizures will occur when possible, but that factors yet to be determined may necessitate that all digital storage devices and media that may reasonably contain the sought after evidence may be seized for offsite review.

The investigator serves the warrant and finds a single computer at Sam's home. The system is on and, according to the suspect, has a Windows XP operating system. Based on the suspect's assertion that the computer is password-protected, and he has not given the password out to anyone, it is reasonable to believe that the computer is used solely by its owner. At this point, the on-scene investigator is staring at a glowing monitor with a happy desktop picture of calming fields and clouds, but the investigator is now faced with a few tough decisions. The computer appears to be running Windows XP, which corroborates the suspect's statement. Windows XP can survive a rapid power loss, so pulling the plug is an option, but pulling the plug means that the entire computer would need to be brought back to the computer forensics laboratory for examination. The investigator knows that the backlog at the computer forensics laboratory is approaching six months—way too long to determine if the suspect is stalking the victim. In six months, the stalking could escalate if there is no police intervention (depending on the type of stalker), and the victim could be physically assaulted. Further, the investigator knows that Windows XP is equipped with the Windows Encrypted File System, a seldom-used folder and file encryption system that, if enabled, would make the recovery of the information on the system very difficult without the suspect's cooperation.

The investigator thinks of other options at his disposal. The investigator could use a software preview tool in an attempt to locate the information stated in the warrant. In this case, Sam uses Microsoft Outlook as his local e-mail client, and a .pst file containing all the Outlook-related folders would exist on the system. This .pst should contain an e-mail in the sent items folder that matches the e-mail received by the victim. If the investigator had reason to believe there was information stored in the RAM that would be relevant to the case, the investigator could dump the RAM for later analysis. This might be the scenario if the investigator notices a draft of another e-mail currently on the screen. If the e-mail is found in the .pst during a preview, the entire drive could be imaged, or just the .pst could be imaged if the investigator has reason to believe that imaging the entire drive would be difficult.

In this example, maybe the investigator would decide to pull-the-plug and deliver it to the lab. Maybe the investigator believes there is enough evidence based on the victim's complaint to have the suspect come to the station for a talk about what is going on. But maybe the investigator's hair on the back of his neck rises up when talking to the suspect and the investigator gets a gut reaction about the level of urgency regarding the case. Maybe the on-scene preview and securing the .pst provides the investigator with enough evidence to take the suspect into custody. The important point is that without additional options to review the digital data, the investigator's hands are tied.

In line with the incremental approach described in the Manual, the investigator may have other options available besides wholesale seizure, such as:

- Previewing information on-scene
- Obtaining information from a running computer
- On-scene seizure of information through the complete imaging of the media
- On-scene seizure of information through the imaging of a specific data object

In the next section, we take a look at the preceding options and discuss how each fits into the larger picture of responding to and investigating crimes with digital evidence.

Previewing Information On-scene to Determine the Presence and Location of Evidentiary Data Objects

The on-scene responder must make conclusions about where the information described in the warrant is most likely to be present on the storage device or media. In the case of a CD or DVD, the preview is much less complicated, as the chances of inadvertently writing to a piece of optical media are much lower than if they were working with magnetic-based media. With a CD or a DVD, the responder could use a forensics laptop running any number of computer forensic tools to quickly acquire and examine the contents of a CD or DVD for review.

A similar process could be conducted for flash-based media, although a greater level of care may need to be taken to ensure the media is not changed. Here, flexibility is once again a critical characteristic. Previewing a few pieces of optical media on-scene may be appropriate, but greater numbers of media may need to be taken off-scene for review at the laboratory.

Technology exists that enables responders to preview the data on the storage media in an effort to locate the information described in the warrant. These "forensic preview software" packages, now in their infancy, are becoming more accepted within the community that investigates crimes involving a computer. The most common preview software packages come on CD and are essentially a Linux operating system that runs completely in the RAM and does not require any resources from the hard drive(s). Several of these disks are in current use by law enforcement, including Knoppix, Helix, and Spada. Several controlled boots will need to be performed to ensure the correct changes are made to the BIOS to direct the computer to boot from the CD. Although best practices should be determined locally, I recommend that the power to all the hard drives in desktop computers be disconnected and that laptop hard drives be removed while controlled boots are conducted to determine how to change the boot sequence in the BIOS. Further information on using controlled boots to examine and change BIOS and CMOS information can be found in the seizure procedures in the publication *Forensic Examination of Digital Evidence: A Guide for Law Enforcement* (NIJ, 2004).

Once the system is booted to the forensic preview software, the computer's hard drives can be mounted, or made available, in Linux as read-only. Once mounted, the preview software will provide the responder with an interface to either search for the desired information through keyword searches, or the responder can navigate through the directory tree in an attempt to locate a given file or directory. If the information described in the warrant is located during a preview, the responder may choose to image the specific data object, file, or folder where the information is located. The responder may also choose to seize the entire hard drive, now that the preview has provided him with a greater level of comfort that this particular "container" includes the desired information.

Over time, these forensic preview software packages will continue to evolve and develop as the problems with wholesale seizure become more evident and the need to focus the seizure of individual data objects from a digital crime scene becomes more apparent. It is hoped that the evolution of these tools will include the addition of features and special characteristics that make a tool "law enforcement specific." The lack of law enforcement specific features, such as intuitive interfaces, audit trail recordkeeping, and the production of evidence-quality data, are often an impediment to the adoption of commercial software by the law enforcement community (ISTS, 2004).

Obtaining Information from a Running Computer

If the investigator encounters a computer that is running, and the investigator believes there is information of evidentiary value stored in the computer's active memory, or RAM, there are options available that allow for the RAM to be recovered. For example, let's examine a

situation where an investigator shows up on-scene at a location where a suspect has been chatting online with a minor or undercover officer. When the officers arrive at the scene, the suspect quickly closes the chat window. By default, many chat programs do not keep a log of the chat sessions and almost all of the actual chat activity happens in a portion of the program running in the computer's RAM. Without being able to obtain a *dump*, or download of the RAM, there would be little chance to obtain any information from the suspect's computer about the chat session that just occurred. Chatting is not the only type of data that would be held in RAM. Passwords, unsaved documents, unsaved drafts of e-mails, IM conversations, and so on could all be held in the RAM, and in no other place on the computer. The investigator needs to make a decision if the information described in the warrant would reasonably be found in the RAM of the computer. If the warrant describes information related to proof of embezzlement, there may be little reason to believe that the data held in the RAM would be relevant to the case. That is not to say that it isn't possible—but the responder needs to go through the process of determining the locations that have the highest probability of containing the information described in the warrant. Even if the suspect had worked on a relevant file and remnants of the same existed in the RAM, it would be logical to conclude that the file would be saved onto more permanent media, such as the hard drive. On the other hand, if the warrant detailed information related to inappropriate chat or instant messaging sessions, the RAM of the running computer would be the primary, and most likely the only, location where the information described in the warrant could exist. In this case, the use of a program such as Helix to "dump" the RAM to the responder's storage device would be a very high priority (Shipley, 2006).

Be careful about what you wish for, however, as the RAM dump could include several gigabytes of semi-random information. Pieces of documents, Registry keys, API calls, and a whole host of other garbage will be interwoven into a gigantic text file. Minimization still is a factor even when the RAM has been identified as being one of the locations where relevant data could exist—if the data might reside elsewhere, it may be more productive to go that route than to attempt to carve it from the RAM dump.

SEARCH, a national law enforcement training organization, recently published a primer on the collection of evidence from a running computer, which involves using preview software to obtain the contents of the RAM from a running machine before seizure (Shipley, 2006). SEARCH's article represents a departure from the norm in that the article recognizes that changes to the computer operating system will occur when a USB drive is inserted into the machine in order to receive the contents of the RAM. However, the important point highlighted by the SEARCH article is that the changes are known, explainable, and do not affect any information that has evidentiary value. "Hold on," you say, "moving the mouse and/or inserting a USB device will change the information on the suspect's drive, and that is strictly forbidden!" In response, I say that there are many in the investigative and legal

communities that see little issue with a law enforcement agent performing operations that changed data on a suspect's hard drive or other media—as long as the agent acted in a reasonable manner and documented their actions appropriately. The firm and absolute stance that data cannot be changed needs to be examined to determine if our cases have been negatively affected by the promulgation of bad advice.

Imaging Information On-Scene

Imaging of an entire hard drive on-scene is fairly common among the more technically savvy digital crime scene responders—even more so for private sector investigators that often face cases where the hard drives need to be examined, but the business in question is not comfortable with letting the original drive out of their possession. In both of these cases, the analysis of the imaged drive usually occurs back at the laboratory. Rarely do you hear of a drive being both imaged and previewed on-scene—although such a process may actually address a number of concerns about the use of preview software to review the information on a drive while on-scene—specifically, performing a preview of the evidence on the original drive.

While the acquisition of an image of a drive on-scene may be fairly common among the more technically skilled, usually for corporate crimes, we find there is little use of this technique by less skilled personnel for low-level crimes. However, there are a number of good reasons to perform imaging on-scene for most computer crimes. First, as mentioned earlier, previews of the evidence can be performed on the imaged copy with less worry about the investigator inadvertently damaging information on the original hard drive. Second, in those instances where outside concerns prevent the seizure of the physical media, such as PPA concerns, third-party data, and multiple users of the computer, the imaging of the hard drive provides another option for the on-scene investigators.

Terminology Alert...

Imaging vs. Copying and Hashes

It is important that the data on the suspect's hard drive be imaged to the destination drive/device rather than just copied. The process of imaging creates a bit-stream copy—or an exact copy of the 1s and 0s—of the information being copied. The regular copy function within the operating system will attempt to write the file according to its logical programming—meaning that the file being written to the drive could be spread across numerous clusters on the target drive. The point of imaging the data is

Continued

that an exact replica of the data as it appears on the source drive is created on the destination drive—specifically the exact order of the bits (the 1s and 0s) on the drive—hence, the term *bit stream copy*. Because imaging preserves the exact order of the bits from the original to the copy, hash functions are able to be run against the entire chunk of the source drive, which is then imaged and compared to the exact replica created on the destination drive. Image hashing allows the responder to mathematically prove that the data that exists on the source drive is exactly the same on the destination drive. Some claim that a few of the hash algorithms (like the MD5 hash algorithm) have been cracked. This is technically true; however, the circumstances for collisions—two different files that generate the same MD5 hash—were specifically created to prove that collisions can occur. The chances of an MD5 hash collision occurring during the comparison of a source drive and an improperly imaged drive would be unbelievably small. I would feel very confident that a hash match between two files/images that are supposed to match to be proof that the two files/images are in fact an exact copy. I feel even stronger about the validity of the next generation of hash algorithms, including SHA1, SHA-256, or SHA-512.

Imaging Finite Data Objects On-Scene

In the current law enforcement climate, there is little discussion of the seizure of particular pieces of information. Generally, the entire computer is seized—and the seized computer is usually called "evidence." The data contained within the computer are reviewed at a later date for any files or other pieces of information that can help prove or disprove a given premise. From an outsider's perspective, it would appear as if the seizure of the entire computer is the preferred method of obtaining the evidentiary information, but we've established that imaging on-scene is fairly well accepted within the digital investigative community. So, are there other options that include the seizure of a finite number of data objects as evidence?

If we can image the entire hard drive on-scene, there is an argument that we can image sections of it. We routinely ask companies and ISPs to do just that when we ask them to preserve evidence of a crime—rarely do we seize the ISP's servers, nor do we ask them to provide an image of the entire server so a computer forensics exam can be performed. Are there reasons why we can't use the same logic when responding to a suspect? The larger question is whether this type of seizure is appropriate. Are there circumstances when a finite amount of information is needed to prove guilt, and the seizure of the original hard drive is not an option? This discussion is very similar to the previous discussion regarding imaging the entire drive on-scene in situations where the physical media cannot be seized. There may also be situations where a finite piece of information would suffice to move the case forward. In these situations, the seizure of a finite number of data objects may be a viable option for responders.

In our case example discussed earlier, where Sam is accused of stalking Sally, let's assume that an arrest warrant hinged on the presence of the harassing e-mail on Sam's computer. If the preview of the computer showed that the e-mail in question existed on Sam's computer, and the investigator had the ability to image the .pst file that contained the e-mail, the investigator could take Sam into custody at this time and have all the evidence needed to wrap up the case. There would be no need to add yet another machine to the computer forensic backlog, and the investigation could be wrapped up immediately, rather than having to wait weeks to months for a completed forensic review.

NOTE

The focus on the seizure of data objects discussed within the other options section does not transfer well to the seizure of computers suspected of containing child pornography. It is strongly recommended that guidance on the seizure of computers containing child pornography be obtained from the Internet Crimes Against Children (ICAC) Task Forces. This network of 46+ law enforcement agencies specializes in the investigation and prosecution of crimes against children facilitated by computer. Additional information about ICAC can be found at www.icactraining.org.

I can hear you yelling "WAIT! What if I think he might have child pornography on his computer?" Good question. If the warrant for the case specifies that the investigator can search for and seize the sent e-mail in question, then it would be hard to justify why the investigator spent all day looking through the suspect's vacation pictures for possible images of child pornography. A warrant for the seizure of a given piece of information that results in the seizure of a computer, or other digital storage device, does not give the law enforcement agent carte blanche to look through every file on the computer. As it relates to the child pornography question, if the investigator believes there is evidence of child pornography on the computer, the investigator is better off obtaining a warrant for the suspected child pornography rather than to search for evidence of one crime under the pretenses of another crime.

That is not to say there aren't instances when you may stumble across evidence of a different crime when reviewing digital information. Should the occasion arise when you are looking for one type of information under a specific warrant, and inadvertently find evidence of another crime, the legal guidance is that you should immediately stop the review and obtain a second warrant to search for evidence of the second crime.

It is theoretically possible that you could finish examining the computer under the first warrant, and not specifically search for items pertaining to the newly discovered crime. However, that strategy is not recommended.

But do we have the tools necessary to enable us to copy-off only the relevant data objects? Can this be done within a reasonable time frame? From a technologist's viewpoint, the technology is often more flexible than the legal framework within which the technology operates. The current technology allows us to search very rapidly through thousands of pages of information for keywords, a feat that would be all but impossible with paper records. But much of the specialized computer forensic tools are designed to be used in a forensic laboratory environment and not for on-scene response. These powerful forensic tools often require a fair amount of time to analyze and process the information on a target drive. Often, these laboratory examinations involve tools that may take hours to complete a given function, and the review of information often involves hours of pouring through documents and graphics. If we consider that "time" is one of the most limiting factors when conducting on-scene analysis, there is definitely a conflict between the best technical analysis that could be performed and the time frame in which a reasonable on-scene analysis should be completed.

The seizure of data objects from large servers while in the course of investigating network intrusion cases is fairly common and accepted, but it is difficult to tell if the seizure of data objects will become more common in the everyday investigator's response toolkit. Although there appears to be a general legal and technological framework within which data object seizure can occur, it is still difficult to swallow the fact that the original evidence will be left behind. The use of this technique on business computers and networks follows the argument that the business is a disinterested third party, and that if relevant data is missed, the investigator can go back and retrieve additional information because the business has no desire to interfere with the investigation. But would a spouse or roommate constitute a disinterested third party with regards to data on their computer? Can we develop tools that give the investigator a greater level of comfort regarding the thoroughness of the on-scene previewing/review? These questions, and others that will spring from discussions like this, will shape the way in which this technique, and the other options presented earlier, become accepted or rejected by the digital evidence response community.

Use of Tools for Digital Evidence Collection

Where the computer forensics of yesterday relied on vary basic tools that allowed manual manipulation of the seized data objects, we have since developed tools that assist in the acquisition, organization, and examination of the data. Both the ubiquity of electronic information and the sheer volume of seized digital information have necessitated the use of tools to assist in the investigative process. Hardware and software write blockers and hard-drive duplication devices have reduced the chances of damaging the information on source drives. Tools beyond simple hex editors and command-line scripts were developed

to assist the examiner in performing keyword searches, sorting data objects by file type and category, and scouring the source disk for file remnants in file slack space and drive free space. Tools like Autopsy Browser, SMART, iLook, Encase, and Forensic Toolkit are dramatic departures from manual command-line searching and have had a significant impact on the efficiency in which large volumes of data are examined. These tools have also increased the accessibility of digital evidence to those outside of the closed circle of highly trained forensic examiners.

The way in which digital information is analyzed has changed over the years—obviously driven by the ever-increasing amount of information stored digitally. But other changes have been driven by the increase in our knowledge of how to work with digital evidence—most notably in the development of tools to assist in different phases of the investigative and forensic process. The use of software and hardware tools by on-scene responders can begin to address how we work toward achieving a greater level of data object seizure. Current tools, such as ImageMasster and Helix, begin to enable an on-scene responder to image an entire drive and to seize the contents of the RAM. Other tools in this domain provide some capacity to preview the contents of a suspect drive and to image only the necessary information, as has been the case for years in the incident response disciplines.

Some will argue that no one should use a tool if they cannot explain exactly what the tool is doing. In the computer forensics realm, this often translates to "no one should use a tool if they cannot perform, by hand, the operations that the tool is performing." There is a fair amount of disagreement on this position. The law enforcement community commonly uses tools where they can explain the basic principal, but not the exact manner in which the tool is accomplishing its task. For example, when an officer is trained on the use of the radar gun, she is taught the principals of the Doppler Effect and how the tool records the very precise timings between the sending of a radar impulse and the receipt of the reflected radar energy. The officer would also be shown how the unit is tested and calibrated to ensure reliability. In this way, the officer understands generally how the tool works—it is not reasonable to instruct them on how to construct the device, nor should the officer be required to manually calculate how the speed of a vehicle is determined from recorded radar signals in order to be a proficient operator of the tool.

That is not to say that we should be able to use any tool without accountability. Tools that are used in the seizure or analysis of digital evidence must be tested. This testing is commonly performed by the organization using the tool—since the tool must be tested within the parameters of the agency's protocols—but larger tool verification efforts are underway at the National Institute for Standards and Technology (NIST). NIST has created tool testing specifications for disk imaging tools, physical and software write blockers, and deleted file recovery programs. A number of products have been tested under this program, and the results look very promising. Almost all of the programs or devices tested actually work as purported. That's not to say there are not issues with the NIST program. Technology

changes faster than the standards development and tool testing processes, and the overall number of standards developed through the NIST program has been, unfortunately, small.

However, placing tools at the disposal of the greater law enforcement community has some significant impacts related to the overall model that we follow when working with digital evidence: If we are able to train officers/investigators on the proper use of a given tool, and the tool has passed muster through testing under a given protocol, whether at their local agency or at the NIST, then the officer/investigator is empowered to take an active role in the recovery of digital evidence and in the investigation on the whole.

It is clear that we do not have all the answers to the technological hurdles worked out, but the technology is often not the limited factor, as was discussed earlier. Understanding that the technology will forever be changing and advancing, the legal community must begin to play an active role in providing the technologists with direction and boundaries. The technologists need to heed the legal guidance, examine how future issues will affect law enforcement, and begin designing tools that will provide a critical edge to the good guys.

Common Threads within Digital Evidence Seizure

The landscape of potential seizure environments is complicated and variations are nearly infinite. The level of knowledge of the on-scene responders includes a wide range of skills and abilities. Because the seizure process will be greatly impacted by the particular hardware and software arrangements and knowledge of the on-scene responder, it is not possible to present one correct way to seize digital evidence, unfortunately. What does exist is a continuum of methods mapped against the complexity of the scene versus the skill of the responders.

There are, however, basic threads that tie any seizure process together. The first thread is that you must be able to explain what steps you took to arrive at a particular destination. It does not matter if you come out of a building with a floppy disk or an entire network, you should be able to replicate each step in the process. If you were presented with an exact replica of the scene, you should be able to refer to your notes and do everything exactly the same from arriving on-scene, to collecting the evidence, to walking out the door. In order to achieve this level of enlightenment, there are two sub-threads: (1) Document everything—and I mean everything. Have one person process the scene while the other one writes down every single, mind-numbing step. The documentation should be as complete as practically possible. If one is working alone in the seizure process, consider using a voice recorder and narrate each step for later transcription. The exact steps taken in the process become doubly important if and when the target computer is manipulated in any way—for instance, moving

the mouse to deactivate the screen-saver, or initiating a shutdown sequence. (2) Confucius is attributed to saying: "To know that you know what you know, and that you do not know what you do not know, that is true knowledge." Translated for relevance for the second sub-thread here, it means that if you don't know what you are doing (or worse, what you just did…), or aren't really comfortable with determining the next steps, *stop*, and revert to a less technical seizure method, or seek assistance from someone more qualified. Your knowledge will be judged by your ability to know what you don't know—when to stop—over the knowledge you do possess.

The second thread is that you should seek the seizure method that best minimizes the digital crime scene. If you can reasonably come up with an "area"—meaning drive, directory, file, and so on—where you believe the evidence will be located, it makes the most sense to look in that specific location for the digital evidence. Limiting or minimizing the crime scene has different implications based on whether the search for digital evidence is occurring on-scene, at the station, or back at the forensic laboratory. On-scene, minimization may include excluding professionally produced and labeled CDs from the seizure. Minimization may also include the use of software tools to preview the contents of a computer for a specific data object. Offsite minimization efforts may include searching only certain keywords or examining only a given file type. Even given our ability to search for and find most anything on a computer, we must remember that not every fact is relevant, and analyses that are 100-percent comprehensive do not exist. At the heart of minimization is the ability to know when to stop while looking for digital evidence.

The third thread is that whatever is seized as having potential evidentiary value must be authenticated by the court before it can be admitted into the case. The ability for the court to authenticate the evidence is a significant issue related to digital evidence. Authentication is governed by the Federal Rules of Evidence Rule 901 (28 U.S.C.), which states "The requirement of authentication or identification as a condition precedent to admissibility is satisfied by evidence sufficient to support a finding that the matter in question is what its proponent claims." The salient point of the definition for our discussions is that digital evidence can be authenticated by providing evidence that shows that it is in fact what it is purported to be. I realize that is a bit of cyclical logic—so let's break down the authentication process further for clarification.

Evidence presented to the court can be authenticated a number of ways, including the identification of distinctive characteristics or by merely what type of evidence it is, as is the case for public records. Evidence may also be authenticated by way of testimony to the fact that the matter in question is what it is claimed to be. Courts have upheld the authentication of documents based on testimony (*U.S. v. Long*, C.A.8 [Minn.] 1988, 857 F.2d 436, habeas corpus denied 928 F.2d 245, certiorari denied 112 S.Ct. 98, 502 U.S. 828, 116 L.Ed.2d 69).

However, in the past, computer forensics has relied less on the testimony of those performing the on-scene seizure and more on the testimony of the computer forensic technician. Where the on-scene responder would be able to testify as to where the hardware was located before seizure, the computer forensic technician would take the position to defend their laboratory techniques. The computer forensics community chose to address the authentication issue by creating exact duplicates of the seized digital information and proving mathematically that the copied information was an exact copy of the seized information—and the courts have supported the position that a duplicate of the information can be submitted in lieu of the original when it can be proved that the duplicate is the same extant as the original (*U.S. v. Stephenson*, C.A.5 [Tex.] 1989, 887 F.2d 57, certiorari denied 110 S.Ct. 1151, 493 U.S. 1086, 107 L.Ed.2d 1054).

As it relates to our options for seizure discussed earlier, there are two salient points for discussion. The first is that the seized data—whether from a RAM dump or as a result of the creation of an image of the drive or file—may be authenticated by the testimony of the investigator that retrieved the evidence from the suspect machine. If the case involved a child pornography photograph, and the investigator saw the photograph during a preview, the investigator may be able to assert that the recovered photograph is the same photograph he saw during a preview. The second point is that the creation and matching of mathematical hashes provides a very high level of proof that the recovered data is an exact copy of the original. Although the best evidence rule states that the original should be provided whenever possible, *U.S. v. Stephenson*, noted earlier, shows that an exact duplicate is satisfactory when circumstances limit the production of the original evidence in court. Hard drives, the most commonly encountered type of storage media, are mechanical devices, and all mechanical devices will fail at some point—perhaps after days, months, or decades—but they will fail. By working off of a copy of the seized drive, and presenting the same in court, the investigator is reducing the chances of completely losing all of the data on the seized drive. Taking steps to reduce the complete loss of the digital information relating to the case is but one of the reasons to justify the use of exact copies over the original data.

The final thread is the admissibility of the evidence. The admissibility of evidence is based on the authentication, and the authentication is based on the proof that the seized object is materially unchanged—proof that can be accomplished by showing a complete chain of custody (*U. S. v. Zink*, C.A.10 [Colo.] 1980, 612 F.2d 511). For digital evidence, the proof that the data is what it purports to be and is unchanged has been accomplished by both testimony and use of the cryptographic hash algorithms. Similar to how the forensic laboratory technician uses the hash function to show that the entire seized drive was copied accurately, the on-scene responder can refer to their detailed notes to testify as to the location of the seized information and show that the hash functions proved that the integrity of the data was not compromised during imaging.

Determining the Most Appropriate Seizure Method

Clearly, there will be cases where the most appropriate action is to seize all the physical hardware at a suspect's location. Perhaps it is the only option that the minimally trained responder has at their disposal. Maybe the forensic preview software didn't support the graphics card for the computer. It's possible that additional keyword searches need to be performed or items need to be carved from drive free space, and both would be better performed in a controlled laboratory environment. There are any number of reasons why the on-scene responder will choose to seize the physical container, and that's ok! The important point is that the most appropriate method of seizure is chosen to match the responder's skill level, and that it appropriately addresses the type of crime.

The minimization stage may provide the investigator with the places—computers, storage media, and so on—that have the highest probability of containing the desired information. A preview on-scene may verify that the information exists. In cases of child pornography possession, the on-scene preview may allow the investigator to take the suspect into custody right at that moment—or at least have some very frank discussions about the material found on the computer. The case may be provided to a prosecutor with just the previewed images, and discussions of sentences and pleas can occur immediately, instead of having to wait for a complete forensics examination. If the case is referred to trial, the full forensic analysis of the seized computer can be conducted at that time. On the other hand, maybe a full examination of the data should be conducted to determine if the suspect has produced any new images of child pornography—information that is critical in determining if an active victimization is occurring and is critical to the overall fight against this type of crime. *This simple scenario shows how the incremental approach and the seizure options discussed earlier are needed so as to even begin to get a foothold on crimes with a cyber component, but that circumstances may force investigators to throw out the incremental approach in favor of a complete examination.*

There are a few other key points relating to physical seizure. The first is that the entire computer will be needed by the laboratory to determine the system time and other settings related to the motherboard. If you plan on only seizing the hard drive, imaging the hard drive on-scene, or only imaging relevant information, follow the methodology outlined by NIJ in the *Forensic Examination of Digital Evidence* (NIJ, 2004) to use controlled boots to record the system time versus a trusted time source.

The second key point is that there are many computers and laptops that do not allow for easy access to the hard drives—which would make any attempts to image on-scene impractical and, as a result, require seizure of the hardware. For example, some laptop designs require the majority of the laptop to be disassembled to gain access to the hard drive. I strongly recommend that the disassembly of laptops or other hardware take place in a controlled laboratory or shop environment—there are just way too many little pieces and

screws, often with unusual head designs, to be attempting a disassembly on-scene. In these cases, the physical seizure of the computer itself may be required even if you came prepared to image on-scene.

The third key point is that there may be other nondigital evidence that could reside with the physical computer. Items such as sticky notes can be found stuck to a monitor; passwords or Web addresses can be written in pencil or marker on the computer enclosure; or items may be taped to the bottom of a keyboard or hidden inside the computer itself. I remember one story of a criminal that hid his marijuana stash inside the computer; the wife asserted that he had child pornography on the computer and the computer examiner—and wife—were amazed when bags of marijuana were found inside the computer enclosure.

One last note: Don't turn off the investigative part of your brain while conducting the seizure. Use all the investigative techniques you learned in the academy and employ during the execution of physical search warrants. You will get much further in the case if you use information from one source (computer/suspect) to gain more information from the other source (suspect/computer)—but remember that Miranda rights may be applicable when having discussions with the suspect.

Summary

There is no doubt that the investigators of tomorrow will be faced with more digital information present in greater numbers and types of devices. Seizing the relevant evidentiary information is, and will continue to be, a critical step in the overall computer forensics process. The current view that the physical hardware is the evidence has now been joined by a different view that the information can be regarded as evidence—whether the hardware or information is viewed as evidence has a dramatic effect on how we "seize" or "collect" evidence both at the scene and in the forensics laboratory.

A number of factors may limit the continued wholesale seizure of the physical hardware. The storage size of the suspect's computer hard drive or storage network may exceed an investigator's ability to take everything back to the forensics laboratory. Full disk encryption, now released as part of the Windows Vista operating system, may foil an investigator's ability to recover any data without the proper encryption key. Further, concerns over commingled and third-party data, covered by the Privacy Protection Act, may impact the ability of an investigator to seize more data than specified in the warrant. Lastly, the increasing amount of seized digital evidence is having an effect on the ability of many of the computer forensics laboratories to complete forensic analyses in a timely manner. Both investigations and prosecutions may be suffering because of delays in the processing of digital evidence.

While the existing seizure methodology is focused on the seizure of hardware, investigators need to be able to select the most appropriate option for seizure according to the situation and their level of technical expertise. There are other seizure options that could be considered by the digital evidence response community. On-site previews using Linux- or Windows-based bootable CDs allow an investigator to review the contents of a suspect's computer in a relatively forensically sound manner. Techniques exist to dump the RAM of a suspect's computer to attempt to recover any information that may be stored in RAM but not written to disk, such as passwords, chat sessions, and unsaved documents. Imaging on-scene is yet another option available to investigators. Full disk imaging—where a complete bit-by-bit copy of a hard drive is created on a black drive—is more common and is currently used by a fair number of investigators. Less common is the imaging of select data objects that have evidentiary value. While still controversial, there appears to be a legal and technological framework that makes the imaging of data objects a viable option.

Clearly, there will always be more digital evidence than we can process within our existing organizational and governmental structures. More trained examiners in the field does not always equate to more trained examiners in the understaffed laboratories nor out in the field. The *time* of the most highly trained personnel is one of our most precious resources. There is no possible way that the limited number of specialists can process electronic evidence at every scene. Not only would they not be able to cover every scene, the laboratory work would undoubtedly suffer. In order to protect the time of the most highly trained and specialized people, those with less technical knowledge need to receive

some level of training that allows them to perform a number of duties normally performed by the specialist. In this way, knowledge and high-technology investigative skills are pushed-down to all levels of responder. That is not to say that *training for first responder* isn't plagued with problems—the knowledge required to properly deploy advanced tools often exceeds the amount of time allotted for such training. We're caught in a Catch-22: all line officers need to be able to seize digital evidence, but the first responder level of training may not fully equip the officers to seize the evidence, and the level of training required to more completely understand the digital evidence seizure process may involve multiple days of training, and multiple days of training on a single topic will most likely not be provided to all line officers.

The level of training will affect the responder's use of technology, and the technology encountered will dictate whether the responder's level of training is appropriate in a given situation. There will be cases where the most appropriate action is to seize all the physical hardware at a suspect location. Perhaps it is the only option that the minimally trained responder has at their disposal, or maybe the technology encountered is so complex that none of the responders know exactly how to handle the seizure.

As it stands now, the forensic collection and analysis system works—sometimes tenuously, and frequently at a snail's pace—however, we will undoubtedly continue to face more change: change coming in the way of new devices, higher levels of inter-connectivity, and the ever-increasing amounts of data storage requiring examination. Will the existing manner in which we go about seizing and examining digital information be sufficient in five years? Ten years? Are there changes we can institute now in the way we address digital evidence that will better position us to face the coming changes?

I hope throughout this chapter that I made myself clear that I am not advocating any one seizure methodology over another—the critical take-away point is that we need to provide our responders with options to choose the appropriate seizure method based on their level of technical skill and the situation at hand. I have found in my work with law enforcement in New Hampshire, and throughout the nation, that crimes that involve a computer closely map to crimes that do not involve a computer—all of it part of the migration of traditional crime into the digital medium. If we expect our law enforcement agents to be responsive to traditional crimes with a high-technology component, we must provide them with the appropriate tools and procedures to enable them to actually investigate and close a case. Asking investigators to send each and every case that involves a computer to a forensic laboratory for review is not a sustainable option. If we don't "push down" technical knowledge to investigators and line officers, the specialists will quickly become overwhelmed and investigations will grind to a halt—a situation that has already begun to occur across the country.

The volume of computer forensic exams is only one factor that is driving us toward changing our approach to digital evidence seizure. As outlined in the previous pages, whole disk encryption, personal data and Privacy Protection Act concerns, and massively large

storage arrays are all playing a part in the move to minimize the amount of information seized from a suspect machine. The landscape is quickly changing, and designing solutions to problems of today will not prepare us for the challenges of tomorrow. It is hoped that the change in focus away from the wholesale seizure of digital storage devices and media, in the appropriate situations, will better prepare our law enforcement agents and private sector investigators for the new technologies and coming legal concerns that the future holds.

Works Cited

Association of Chief Police Officers and National High Tech Crime Unit. 2004. *Good Practice Guide for Computer based Electronic Evidence*, Version 3.0. Available on the Internet at http://www.acpo.police.uk/asp/policies/Data/gpg_computer_based_evidence_v3.pdf (12/2006).

Bloombecker, Buck. *Spectacular Computer Crimes: What They Are and How They Cost American Business Half a Billion Dollars a Year*. 1990. Homewood, IL: Dow-Jones Irwin.

Carrier, B. and E. Spafford. "Getting Physical with the Digital Investigation Process." *International Journal of Digital Evidence*. Volume 2, Issue 2, 2003. Available at http://www.ijde.org (12/2006).

Computer Crime and Intellectual Property Section (CCIPS), Criminal Division. "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations." United States Department of Justice. Washington, DC. 2002.

Gilder, G. "The Information Factories." *Wired Magazine*. Volume 14, Number 10, 2006.

ISTS. "Law Enforcement Tools and Technologies for Investigating Cyber Attacks: Gap Analysis Report." Institute for Security Technology Studies, Dartmouth College. Hanover, NH. 2004.

ISTS. "Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Research and Development Agenda." Institute for Security Technology Studies, Dartmouth College. Hanover, NH. 2004.

Meyers, M. and Rogers, M. "Computer Forensics: The Need for Standardization and Certification." *International Journal of Digital Evidence*. Volume 3, Issue 2, 2004. Available at http://www.ijde.org (12/2006).

Moore, Robert. *Cybercrime: Investigating High-Technology Computer Crime*. Anderson Publishing, LexisNexis Group. 2005.

National Institute of Justice (NIJ). *Forensic Examination of Digital Evidence: A Guide for Law Enforcement*. Office of Justice Programs, U.S. Department of Justice, Washington, DC. 2004.

National Institute of Justice. *Electronic Crime Scene Investigation: A Guide for First Responders*. Office of Justice Programs. U.S. Department of Justice. NIJ Guide Series. Washington, DC. 2001.

National Security Agency Information Assurance Solutions Technical Directors. *Information Assurance Technical Framework*, Release 3.1. 2002. Available at http://www.iatf.net/framework_docs/version-3_1/index.cfm.

Nolan, Joseph R. and Jacqueline Nolan-Haley. *Black's Law Dictionary*, 6th ed. St. Paul, MN: West Publishing Company. 1990.

School of Information Management Systems (SIMS). "How Much Information?" University of California Berkeley. 2003. Available on the Internet at http://www2.sims.berkeley.edu/research/projects/how-much-info-2003.

Shipley, T. and H. Reeve. *Collecting Evidence from a Running Computer: A Technical and Legal Primer for the Justice Community*. SEARCH, The National Consortium for Justice Information and Statistics. Sacramento, CA. 2006. Available on the Internet at http://www.search.org/files/pdf/CollectEvidenceRunComputer.pdf (12/06).

"Scientific Working Group on Digital Evidence (SWGDE) and International Organization on Digital Evidence. Digital Evidence Standards and Principals." *Forensic Science Communications*. Volume 2, Number 2, 2000. Federal Bureau of Investigation. U.S. Department of Justice. Washington, DC.

Sterling, Bruce. "Hacker Crackdown." *Project Guttenburg*. Champaign, IL. 1992. Available on the Web at www.gutenberg.org/etext/101.

Technical Working Group for Electronic Crime Scene Investigation, Office of Justice Programs. *Electronic Crime Scene Investigation: A Guide for First Responders*. U.S. Department of Justice, National Institute of Justice. NIJ Guide series, NCJ 187736. Washington, DC. 2001.

United States Secret Service (USSS). "Best Practices for Seizing Electronic Evidence." 2006. Available on the Internet at http://www.secretservice.gov/electronic_evidence.shtml (12/2006).

United States Department of Justice. *Federal Guidelines for Searching and Seizing Computers*. United States Department of Justice. Washington, DC. 1994.

Federal Rules of Evidence (FRE) are available at judiciary.house.gov/media/pdfs/printers/108th/evid2004.pdf.

Federal Rules of Criminal Procedure (FRCP) are available at judiciary.house.gov/media/pdfs/printers/108th/crim2004.pdf.

Additional Relevant Resources

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 US, 579 (1993).

Noblett, M., M. Pollit, and L. Presley. "Recovering and Examining Computer Forensic Evidence." *October Forensic Science Communications*. Volume 2, Number 4, 2000. Federal Bureau of Investigation. U.S. Department of Justice. Washington, DC.

Duerr, T., N. Beser, and G. Staisiunas. "Information Assurance Applied to Authentication of Digital Evidence." *Forensic Science Communications*. Volume 6, Number 4, 2004. Federal Bureau of Investigation. U.S. Department of Justice. Washington, DC.

Brown, C. and E. Kenneally. "Risk Sensitive Digital Evidence Collection." *Digital Investigation*. Volume 2, Issue 2, 2005. Elsevier Ltd. Available on the Internet at http://www.sciencedirect.com/science/journal/17422876.

Brenner, S.W. and B.A. Frederiksen. "Computer Searches and Seizures: Some Unresolved Issues." *Michigan Telecommunications Technical Law Review*. Volume 8, Number 39, 2002.

Joint Administrative Office/Department of Justice Working Group on Electronic Technology in the Criminal Justice System. "Report and Recommendations." 2003. Available on the Internet at http://www.fjc.gov/public/pdf.nsf/lookup/CompInDr.pdf/$file/CompInDr.pdf (12/06).

Wright, T. *The Field Guide for Investigating Computer Crime: Parts 1–8*. 2000–2001. Available on the Internet at http://www.securityfocus.com/infocus/1244 (12/2006).

Frequently Asked Questions

Q: What is your opinion on the certification of personnel? Can't we fix all the problems regarding experts and admissibility of evidence once personnel are certified?

A: Certification of personnel is, in my opinion, counterproductive. One of the more commonly seen certifications is vendor certification. These trainings are generally useful as long as the training certified that they attended training, not that they are certified in the use of a tool. Another option is to obtain a certification through an independent certifying body. A number of these types of organizations exist and they do provide a means by which people can advertise their level of knowledge and skill, which is rather handy when reaching out for assistance across jurisdictional boundaries, as often occurs while investigating crimes with a cyber component. However, it is highly unlikely that the court system will give carte-blanche acceptance to a particular certification. If you were to testify as an expert, your certifications may assist you in passing muster as an expert witness, but the certification won't be an automatic bye onto the stand.

Some last thoughts on certifications: Let's assume for a minute that Congress took up this issue and passed a law requiring that all computer forensic examiners must be a Certified Forensics Guru. As soon as the first person achieves the certification, it means that everyone else, by default, is not certified. Forensics personnel would need to spend time working on obtaining the certification, time that should be spent on existing cases. Finally, how would such an overarching certification affect onsite acquisition, live-forensic previews, and the seizure of digital evidence? Although there may be some benefits to such a certification, the negatives, particularly related to empowering all law enforcement to play a role in investigating crimes with a cyber component, appear to outweigh the potential positive affects.

Q: Is the seizure of data objects or evidence preview relevant when a computer or other device is actually stolen?

A: In the instance where the digital device was actually stolen, or generally when the hardware or media represent the instrumentality or fruits of a crime, then it is again appropriate, without question, to seize the physical hardware or media. In these cases, the hardware or storage media may itself be the "evidence" and *there may not necessarily be a need to examine data objects on the computer or device* (CCIPS, 2001). These types of seizures show why it is important to understand exactly how the computer was used in committing the criminal act. It is important to remember that not all crimes that involve a computer will necessarily involve digital evidence. What is worse is that many of these seized devices are needlessly processed by an overtaxed computer forensic system. As discussed earlier, remember to keep computers and digital devices in perspective, and look to use digital evidence only when appropriate.

Chapter 9

Conducting Cyber Investigations

Solutions in the chapter:

- **Demystifying Computer/Cyber Crime**
- **Understanding IP Addresses**
- **The Explosion of Networking**
- **The Explosion of Wireless Networks**
- **Interpersonal Communication**

☑ **Summary**

☑ **Frequently Asked Questions**

Introduction

We often fear most what we don't understand. That could be said about computers and the investigation of computer crimes. Many investigators cringe at the mention of a computer and seek to offload any computer-related crime to the "computer crime guy" in their office. Although computers have been around for a few decades, they've finally reached levels where it is feasible to expect that everyone has access to a computer. The computer is no longer a "nice to have," it is a "must have." Those who don't own their own computers can walk into a public library or cyber cafe to gain access to a computer. Similarly, access to the Internet is becoming ubiquitous through connections provided by libraries, coffee shops, computer stores, and even fast food restaurants. This explosion of computer technology and acceptance has opened up a whole new world of opportunity to the criminal element that constantly looks for new ways to exploit people through time proven scams and tactics. As computers become more deeply integrated within society, it is likely that a computer or similar type device will play a role in criminal activity. A basic understanding of computers is all that investigators will need to learn that *computer crime* is just plain old crime packaged up in a shiny new wrapper.

Demystifying Computer/Cyber Crime

Computers start to play a role in crime in situations where the capabilities of the computer allow a person to commit that crime or store information related to the crime. An e-mail phishing scam is a common example where the bad guy generates a fictitious e-mail for the sole purpose of enticing people to a spoofed site where they are conned into entering sensitive personal information. That sensitive information is then available to the bad guy in order to perpetrate an Identity Theft. In another example, a suspect might use the computer to scan and generate fake bank checks, or create fake identification. In both of these cases the crime required the inherent capabilities of the computer for its commission.



WARNING

The mere presence of a computer does not make a crime a computer crime. We must be careful not to hastily label a crime a "computer crime" just because a computer was involved. What if the new laptop I purchased was stolen from my vehicle while I was in the convenience store getting milk? This would not be a computer crime just because a computer was involved, but a theft. How about an office fight where an employee strikes another with the keyboard of their computer—should we call out the Forensic team? Absolutely not (well, maybe, if the assault resulted in a homicide). The computer in and of itself is not important, it is just merely an object like many others in our lives.

Since computers are so pervasive, it is an absolute necessity that investigators learn how to investigate crimes that involve a computer. The basic design of computers—including vast amounts of storage and meticulous file timestamping—can make them a wealth of evidence as traces of the crime can often be retrieved by an experienced investigator. This does not mean that every investigator needs to become an expert in computer technology, but there are basic concepts and methods that must be learned in order to develop old school leads. *The key is to gain at least some basic computer knowledge and skills to put you ahead of the average computer user; skills that allow you to apply traditional policing skills and procedures to the case.*

The crimes that are being committed haven't changed, just the manner in which they're being committed. Think about it. Back before the Internet, the telephone, the telegraph, and the Pony Express, if a person wanted to threaten to kill someone, it was likely they would have to physically place themselves in proximity to the person and speak that threat. As services and technologies developed, new ways emerged through which a person could commit that same threatening act. They could send a letter, a telegram, or even better, make a phone call. Now we can send an e-mail or instant message (IM). Same crime; same underlying elements and facts to be proven. They only change is the manner of delivery. The key to a successful investigation of a computer crime is the development and follow-up of case leads. Although many leads will dead end, it is the one that continues to develop into further leads that can end up solving your case. Many believe that investigations involving computers are above their capabilities, but that is often not the case. By learning and adapting some basic computer knowledge and skills, today's investigator can react to new technologies and still develop workable old school leads.

NOTE

IM stands for instant message. Instant messaging is another way for people to communicate with each other by computer in real time. A chat session is established between two or more computers using compatible applications through which written messages and files can be transmitted back and forth. The unique challenge of instant messages is that their content is not often recorded by service providers or the applications facilitating the chat. Once the IM session closes, the contents tend to be lost. This is not always the case as users can turn on chat logging, but by default most chat applications do not record sessions.

Throughout this chapter, critical skills will be discussed that prepare an investigator to deal with computer crime investigations. By developing a basic understanding of key concepts and learning to apply basic computer skills, an investigator can learn how to proceed with computer crime cases much in the same way as traditional cases. Issues such as IP Addresses, Networks, Wireless Devices, and Interpersonal Communication will be discussed

with the sole purpose of providing the investigator with a basic understanding of each topic area and the skills that can be employed to yield workable physical leads. Many of these skills will build the foundation of computer crime investigations not only today, but well into the future as these technologies expand and become more complex.

Notes from the Underground…

"Application Stupid"

Even though computers have been in our society for quite some time, it is still arguable that many within the population are not highly skilled with them. With the prevalence of computers today, it becomes increasingly important for computer and software companies to develop systems and applications that are "user friendly." These user friendly devices and systems are intended to make people's lives easier. People, being creatures of habit, are often quick to embrace any solution that will allow them to work less. This has facilitated the rapid acceptance and integration of complicated systems into everyday life.

In the quest to make applications and operating systems easier for the end user, programmers have had to develop very advanced and complicated programs. There is a direct correlation between the ease of use by the end user compared to the complexity of the underlying code that is required for the application to run. Many operating systems today are so advanced compared to their earlier versions that little interaction is required of the end user to install new programs or add peripherals. The system itself is able to identify new devices or programs, load the necessary supporting drivers, and set parameters to make the new device or program function. All this is done for the benefit of the end user, who is no longer required to have a fundamental understanding about how the computer and/or its software functions.

A large majority of people are what I call "application stupid"; the process of using a computer or application is so simplified that the user is not required to possess any enhanced level of computer skill or knowledge. The user is able to operate the computer, sometimes at a fairly high level, without having any understanding about what is going on in the background. Application stupidity can provide an opportunity to the investigator to obtain traces of information or evidence that has been left behind as a result of the complexity of the application or operating system versus the rudimentary skill of the user. For example, a suspect creates a file on their computer that is incriminating in nature, they delete it, and then they empty their recycle bin. They believe that the file no longer exists—which is not the case. Their limited knowledge about how operating systems handle deleted files has created an opportunity for the investigator to retrieve the deleted file.

Continued

The simpler the program is to the end user, the more complex the coding; the more complex the coding, the more likely that fragments of information will be left behind. The theory of application stupidity is likely to become more pervasive as the complexity of operating systems and programs increase to keep pace with a growing user base that demands simplicity.

Understanding IP Addresses

All law enforcement investigators need to understand the basics of IP addressing in order to trace users of the Internet to a physical location. Much like a phone number that shows up on a caller id box from a threatening phone call can provide an investigator with a specific starting location for their investigation, an IP address can provide that same type of lead. By understanding what IP addresses are, how they're assigned, and who has control over them, an investigator can develop workable case leads.

IP addresses provide a connection point through which communication can occur between two computers. Without getting into too much detail about them, it is important that you understand how to identify an IP address when you see one. These addresses are made up of four 8-bit numbers divided by a ".", much like this one: 155.212.56.73. Currently the Internet operates under the IPv4 (Internet Protocol Version 4) standard. In IPv4 there are approximately 4 billion IP addresses available for use over the Internet. That number will be expanding in the near future to about 16 billion times that number when transition is made to IPv6.

During the birth and initial development of today's Internet, IP addresses primarily were assigned to computers in order for them to pass network traffic over the Internet. Computers were physically very large, extremely expensive, and pretty much limited to the organizations that controlled the primary networks that were part of the primordial Internet. During this time, an IP address most likely could be traced back to a specific computer. There are a limited number of large organizations that own and control most of the IP Addresses available with IPv4. Therefore, if an investigator has been able to ascertain the IP address of an illegal communication, they will also be able to determine which organization owns the network space within which that address is contained. That information in and of itself will often not be enough since many of these organizations sublease blocks of the IP Addresses they own to smaller companies, such as Internet Service Providers (ISP). It will be the investigative follow-up with the ISP that is likely to provide the best results. Using an analogy, we can think about IP addresses much like phone numbers, where the major corporations are states and ISPs are towns or calling districts. If an investigator was following up on a case involving a phone number, the area code would narrow the search down only to a particular state, and the remaining numbers would identify a particular account.

Remember that for Internet traffic to occur, an external IP address must be available to the device. Access to an external IP address is provided by an ISP. ISPs sublease blocks of IP addresses from one or more of the larger corporations that control address space and in return

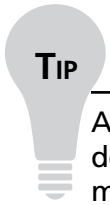
they will in essence sublease one of those addresses to the individual customer. This connection to the Internet is most often done through a modem. Modems came in varying configurations such as dial up, cable, and DSL. Depending on at what point in time you began using the Internet, you may already be familiar with these devices. The older of the three listed is the dial-up modem that required the use of a telephone line. When users wanted to connect to the Internet, they would plug the modem installed in their computer to their phone line and then dial one of the access numbers provided by the ISP. The dial-up modem is the slowest of the available devices that can make the transfer of large files a painful process. Therefore when dealing with cases that require large file transfers such as child pornography, it is less likely that a dial-up connection would be used. A distinct advantage of the dial-up modem, though, is the portability since the connection can be made on any phone line by dialing an appropriate access number and providing valid account information.

More common today is Internet service provided through TV cable or through DSL (Digital Subscriber Line); both of these services provide higher connection speeds making the transfer of large files relatively easy. When a consumer contacts an ISP about Internet access, typically they are assigned an installation date when a technician comes to the residence to connect the necessary wiring to the home through either their cable provider (cable modem) or phone provider (DSL). With the appropriate wiring in place, an external modem is connected to the line provided through which the computer in the home will connect. The modem provides the interface through which the home computer can be physically connected to the Internet.

When the home user is connected to the ISP's physical connection to the Internet, the ISP must still assign the home user's computer an IP address in order for the computer to communicate over the Internet. IP addresses are assigned two ways, statically and dynamically. If static addressing was to be used, the install technician would configure the computer's network interface card (NIC) with the specific IP address during install. Static assignment by an ISP would limit the total number of customers an ISP could have by the total number of external addresses they control. Let's say that XYZ ISP had subleased a block of IP addresses from a large corporation in the amount of 1,000 unique valid addresses. If that ISP statically assigned addresses to their customers, then the total number of customers they could have on the Internet would be limited to 1,000. Leasing blocks of external IP addresses is very expensive as the demand is high compared to availability. ISPs realize that it is unlikely that all their customers will be on the Internet at the same time, so in order get the largest return on their investment, they use an addressing scheme called dynamic addressing, which allows for computers that are actively connected to the Internet to be assigned an unused IP address.

Here's how dynamic addressing works. XYZ ISP has 1,000 addresses available to their customers. They set up a server, referred to as DHCP server, which maintains a list of the available addresses. At installation, the technician sets the consumer's computer NIC to get an address assignment through DHCP. When the consumer's computer is turned on and connected to the network, the NIC puts out a broadcast requesting an IP address assignment. The DHCP server responsible for the assignment responds to the request by providing

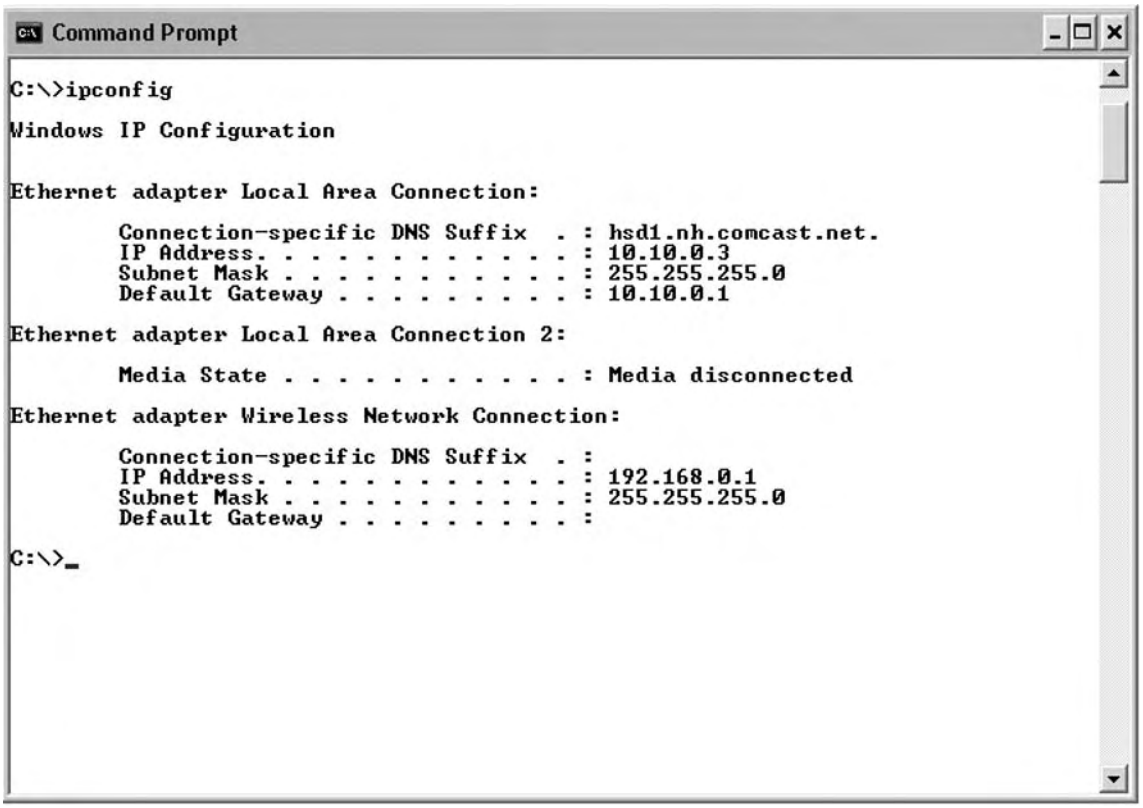
an IP address from the pool of available addresses to the computer's NIC. The length of time that the computer will use that assigned address is based upon the "lease" time set by the DHCP server. Remember that the ISP wants to have the maximum number of customers using the smallest number of addresses, so the ISP wants to ensure that any unused addresses are made available to other computers. The lease time determines how long that address will used before the NIC will be required to send out another broadcast for an IP address. The IP address returned after the reassignment could be the same address used previously or an entirely new address, depending on what's available in the server pool.



TIP

A number of details about the configuration of a computer's NIC(s) can be determined in Windows by using the ipconfig command at the computers command prompt—most importantly the computer's IP Address (see Figure 9.1).

Figure 9.1 ipconfig Command


Command Prompt
C:\>ipconfig
Windows IP Configuration
Ethernet adapter Local Area Connection:
Connection-specific DNS Suffix . : hsd1.nh.comcast.net.
IP Address. . . . . : 10.10.0.3
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . : 10.10.0.1
Ethernet adapter Local Area Connection 2:
Media State . . . . . : Media disconnected
Ethernet adapter Wireless Network Connection:
Connection-specific DNS Suffix . :
IP Address. . . . . : 192.168.0.1
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . :
C:\>

Note that this example provides details on three different NICs; two physical Ethernet ports identified by the Local Area Connection designation and one wireless network connection. Each NIC can possess a different IP address. IP addresses are important because each device that communicates over the Internet must have an address. In a computer crime investigation involving the Internet, it is very likely that the investigator will need to track an IP address to a location—preferably a person. As discussed earlier, ISPs control the assignment of IP addresses, and ISPs can provide the link between the IP address and the account holder. Understanding the distinction between static and dynamic IP assignment is very important because the investigator must record the date/time that IP address was captured. If the ISP uses DHCP, the IP address assignments can change—investigators need to be sure that the account holder identified by the ISP was actually assigned the IP address in question when the illicit activity occurred.

Let's take a moment and think about this. You're investigating an e-mail-based criminal threatening case where you were able to determine the originating IP address of the illegal communication. You were able to determine which ISP controls the address space that includes the IP address in question. If ISPs use dynamic addressing, how are you going to be able to determine which subscriber account used that address if any of a thousand or more could have been assigned to the suspect's computer? In this case, it would be extremely important for you to also record and note the date and time of the originating communication. The date/time stamp can be matched against the logs for the DHCP server to determine which subscriber account was assigned the IP address in question at that time.

The Explosion of Networking

Much like ISPs use dynamic addressing to maximize the number of customers they could have using a limited number of addresses, customers began using routers to increase the number of computers they could use in their homes that could share that IP address provided by the ISP. The router passes network traffic back and forth between the Internet and all the home computers in the residence connected to that network router. All the network traffic sent from the home computers through the router to the Internet will be seen as coming from a single IP address. The investigator who traces an IP address back to a router will need to do more case follow-up at the location to determine if there is more than one possible computer involved. Analysis of the router configuration and/or logs may provide more information about the computer requesting and receiving the illegal traffic as information, such as the computer's hostname, internal IP address, or MAC address.

Networks have become common place today as the cost and implementation of computer systems has dropped dramatically. Years ago, computer systems were very large (room size) and extremely expensive. This limited the organizations that could afford to use computers in any meaningful way. Today, computers are much more powerful and affordable. This has allowed both companies and individuals to purchase and use numerous computer

systems to accomplish specific needs. The concept of networks, much like the Internet, allows multiple computers to become interconnected to each other in order to share files and resources. The computers on the network will still need to be assigned IP addresses in order to communicate with other computers on the network—but the addresses assigned within a network behind a router, *or gateway*, will fall into the category of internal IP addresses. Unlike the external address assignments required to send and received information on the Internet, internal IP addresses allow computers within a network to communicate with one another. In order for computers on these private networks to access the Internet, there is likely to be an established gateway that has been assigned a single external IP address to be used by all computers on the network.

NOTE

Internal IP addresses can also be used to set up more than one computer into a network environment. When computers are placed within a network, they will be able to see the existence of each other on the network and can be used to pass communications and share files. This is completely independent and not reliant upon having access to the Internet. However, without some type of Internet access, the communications transmitted over that internal network (most often referred to as a private network) would remain within that network and would not be accessible to other computers not physically included in its scheme. Private networks are very common in corporate environments where large numbers of employees need to access or share files with other employees, but for security purposes, no Internet connection is included in order to stop possible unauthorized access from outside the network.

Gateways become a transfer agent for computer traffic between computers on the network and the Internet. This means that the network owner is only required to assign a single external IP address to the gateway in order for one or hundreds of the computers on the network to access the Internet. This provides a challenge to investigators who have been able to trace back that IP address to the gateway owner. The IP address no longer identifies a specific computer directly, but merely identifies the gateway that handed the traffic on to the Internet on behalf of all the computers on the private network. More follow-up must be performed in order to establish the identity of the system that sent the request to the gateway initially.

A benefit of investigating a traditional wired network is that the number of devices connected often is limited to the location at hand and physical limitations of transmission over wired lines. Being able to trace an IP address back to a particular location and network

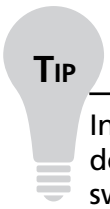
greatly helps reduce the total overall number of suspects. If other identifying information such as the internal IP address, hostname, and MAC address has been determined, then the ability to narrow the suspect down to a single device is greatly increased. If the device is found, then traditional investigatory techniques can be used.

Hostname

Hostnames are the system names assigned to a computer by the system user or owner. These names are used to identify a computer in a network in a format that is easiest to understand by people. If there are multiple computers in the network, each could be given unique identifying names making them more easily recognizable, such as Receptionist PC or Dave's Laptop. The naming choice selected might help to identify the likely location or user of that system. If for example you were investigating a threatening e-mail that had originated from a computer within a network named "Jedi," you might look for people who have access to the network who are also fans of the *Star Wars* series. Keeping in mind that the names can be changed by the user at any time, the matching or nonmatching of a hostname to a suspicious communication or activity is by no means conclusive in itself.

MAC Address

MAC addresses are the identifying number assignment given to NICs that provide network connectivity. That connectivity can be wired or wireless depending on the type of NIC present. MAC addresses are unique to every NIC and would be most equivalent to a serial number. This means that if an investigator is able to determine the MAC address of the device used in the crime, then the device containing the NIC could be identified specifically. However, just like a hostname can be changed, MAC addresses can also be changed through a process called MAC spoofing. Whether or not a MAC address matches a particular communication is not in itself conclusive evidence that the computer containing the NIC was or was not responsible.

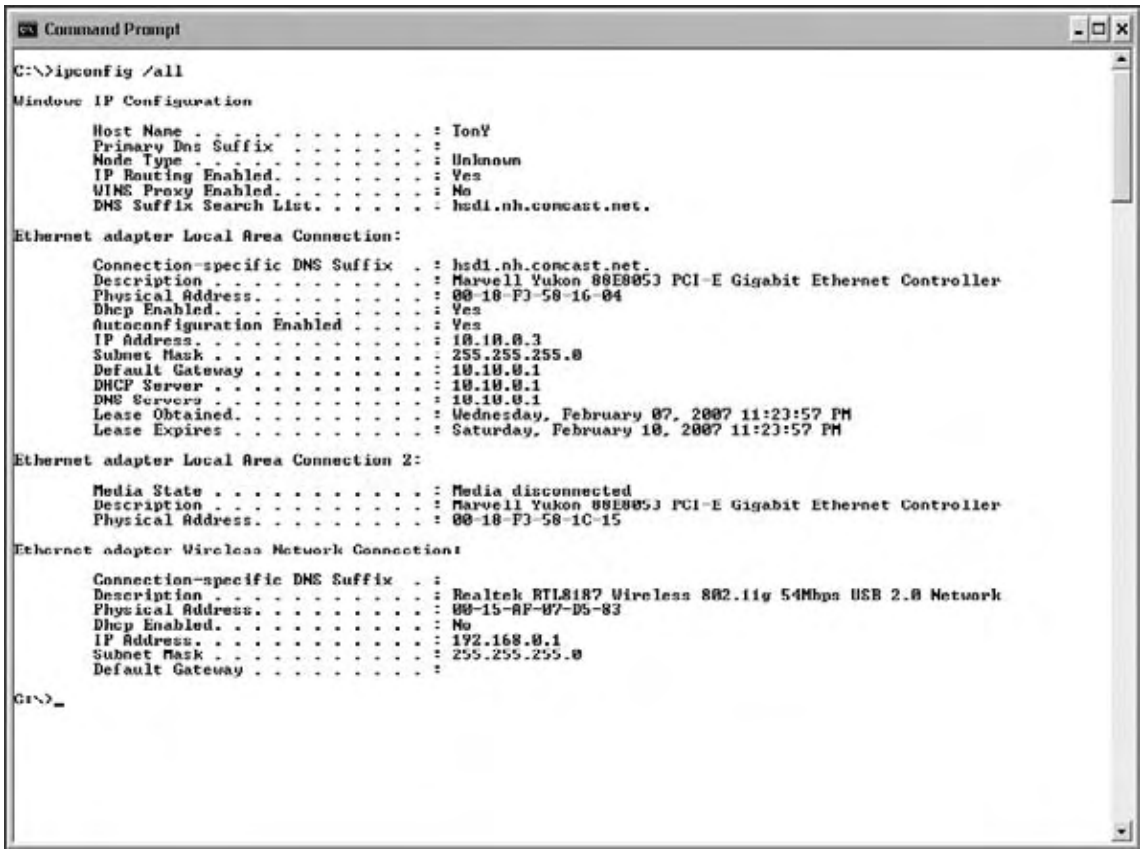


TIP

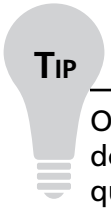
In the previous Tip we learned that the ipconfig command can provide some details about a computer's network interface card configuration. There is a switch that can be added to the ipconfig command that provides more detail about the NIC configuration. At the command prompt, ipconfig /all is used (see Figure 9.2).

You will notice that other details have been provided that are not seen in the ipconfig command, including the computer's hostname, and each of the NICs' MAC addresses.

Figure 9.2 ipconfig/all **Command**


Command Prompt
C:\>ipconfig /all
Windows IP Configuration
Host Name . . . . . : TonY
Primary Dns Suffix . . . . . :
Node Type . . . . . : Unknown
IP Routing Enabled. . . . . : Yes
WINS Proxy Enabled. . . . . : No
DNS Suffix Search List. . . . . : hsd1.nh.comcast.net.
Ethernet adapter Local Area Connection:
Connection-specific DNS Suffix . : hsd1.nh.comcast.net.
Description . . . . . : Marvell Yukon 88E8053 PCI-E Gigabit Ethernet Controller
Physical Address. . . . . : 00-18-F3-58-16-04
Dhcp Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
IP Address. . . . . : 10.10.0.3
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . : 10.10.0.1
DHCP Server . . . . . : 10.10.0.1
DNS Servers . . . . . : 10.10.0.1
Lease Obtained. . . . . : Wednesday, February 07, 2007 11:23:57 PM
Lease Expires . . . . . : Saturday, February 10, 2007 11:23:57 PM
Ethernet adapter Local Area Connection 2:
Media State . . . . . : Media disconnected
Description . . . . . : Marvell Yukon 88E8053 PCI-E Gigabit Ethernet Controller
Physical Address. . . . . : 00-18-F3-58-1C-15
Ethernet adapter Wireless Network Connection:
Connection-specific DNS Suffix . :
Description . . . . . : Realtek RTL8187 Wireless 802.11g 54Mbps USB 2.0 Network
Physical Address. . . . . : 00-15-AF-07-D5-83
Dhcp Enabled. . . . . : No
IP Address. . . . . : 192.168.0.1
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . :
C:\>

Being able to determine the computer's MAC address is a useful skill for investigators. At one organization, network security had set alerts in their system to notify the system administrator when MAC addresses of stolen equipment appeared on the network. The systems administrator notified law enforcement that a stolen laptop had just connected to one of the organization's wireless access points, and they were able to direct the officer to the general area in range of the given access point. The officer was able to make a directed patrol of the area looking for anyone using a laptop that matched the general description of the stolen laptop. Unfortunately the officer was not aware of the ipconfig/all command. Knowing that command would have allowed the officer to conduct field interviews and request consented permission to check the MAC address of any of the suspected laptops against the recorded MAC address of the stolen laptop.



TIP

Once investigators have narrowed the scope of their network investigation down to one computer, they may want to consider the following lines of questioning:

Who has access to the device?
Did they have access on the date and time in question?
Did they have motive?
Is there evidence still on the device that can be retrieved?
What information does the suspect provide?

The Explosion of Wireless Networks

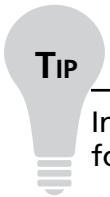
In the not too distant past, networks were isolated to corporate and government entities using large computer clusters and a wired infrastructure. It was less common to find homes with a computer; much less a network. All of that changed with the advent of wireless technology. Many homes and consumer establishments contain private and/or open networks providing access to the Internet, network devices, or offline storage. Cellular companies also compete within the wireless space and offer numerous Internet-enabled devices that allow consumers to stay connected. This proliferation of interconnected and overlapping wireless networks allows criminals to be more portable, creating a heightened challenge to law enforcement to first locate the origin of the action or communication.

Hotspots

Hotspots refer to locations where wireless Internet services are readily available to any user. Some are fee-based and others are offered as a free service to attract customers. In the fee-based system, the person connecting to the network is required to submit valid payment information prior to being granted access. As a service to attract customers into their establishments, many businesses now offer free Internet. This means that anybody entering the establishment, or within range of their wireless signal, can utilize their network to gain access to the Internet. These free hotspots can pose a significant problem for law enforcement since an IP address traced back to any establishment that is set up as a free network is likely to leave the investigator with a large suspect pool—basically anyone within range of the network.

In these situations, the timestamp of the illegal or suspicious activity continues to be critical to the investigation. Knowing the date and time of the alleged incident would allow you to narrow down the pool of possible suspects. A pattern of illegal activity from the address might help build a profile of the offender sufficient enough to jog the memory of employees about a "regular" who visits the location during those time frames. Of course, be careful not to exclude employees in the pool of possibilities unless they can be eliminated based on work

assignments and schedule. Tracing back the IP address will provide only a lead toward where the investigator should look further. It will be traditional investigative skills that will help yield a possible suspect. Understanding IP addresses, hostnames, and MAC address assignments will be crucial to matching your suspect's device to the router configuration and/or traffic logs.



TIP

Investigators working cases involving wireless networks should consider the following lines of inquiry:

Do the employees remember anything unusual during those time periods?

Is the establishment equipped with video cameras and is there footage of the time period in question?

Does the investigator have a possible suspect photo, sketch, or other information that might help in the follow-up?

Does the router providing the service maintain activity logs? If so, what was the computer name and MAC address of the device that perpetrated the activity in question?

Wardriving

As people learn to appreciate and utilize new technologies, they can inadvertently open themselves up to an opportunist who prays on that innocent lack of understanding. Wireless technology is a perfect example. People have longed for the day when they wouldn't be forced to sit at the same desk or location in their home or office to use a computer, but could move about freely without the constraint of wires. Laptop devices have evolved to the point that they are lighter, more portable than, and just as efficient as full-size desktop computers. Most now come equipped with a wireless card as a standard device, which means that the only new device needed to achieve true portability at home or office is the installation of a wireless router.

Wireless routers are so inexpensive and easy to set up that many homes and offices are now wireless enabled. Many wireless routers come with installation CDs that automate the entire process using default settings that will work with most devices. This means that within a few quick steps of returning home with this device, the average person can have a fully functional wireless network established that will communicate with the wireless-enabled laptop they already own. Using the old adage "if it works don't fix it." many will make no attempt to secure that network from outside intrusion. They will not be aware that they have just created an open wireless network that is available to anyone within the wireless signal range.

There are those that drive through neighborhoods looking for the presence of open, unsecured wireless networks. This process is referred to as *wardriving*, and requires no special equipment other than a wireless-enabled device that is capable of detecting wireless signals.

Some will record the location of these networks for their own personal use, still others might post the locations on the Internet as part of a greater hotspot map for anyone to utilize. They types of crimes that can be perpetrated using one of these locations varies. First, the intruder may use the network only as free access to the Internet with no illegal intent other than nonpermissible use of that network and Internet account.

Some people may use this opportunity to scan the network, looking for devices within the network that have known vulnerabilities that they might be able to exploit in order get account and password information. This network could be used to send threatening e-mails, launch viruses, or transfer child pornography. An investigator who has been able to trace the IP address back to the home owner account would need to use some traditional policing skills, which might include interviewing residents, consented or warranted searches of Internet-enabled devices, and review of the wireless router's configuration and log files. Computer skills will lead the investigator to the location, but traditional police work will tie everything together.

Security Alert…

Investigating Wireless Networks

There are situations where a homeowner may contact an investigator about unauthorized access of his or her wireless network. Since most routers have the ability for logging and e-mail alerts about certain activity, an investigator with consent of the network owner could set the configuration of the system to generate log files and e-mail the investigator when suspicious activity is occurring. Remember that in order for a person to use the network, he or she would have to be within range of the signal. If an investigator knew the activity was occurring in real-time, he or she might be able to locate a suspect based on activity in the neighborhood. What other houses appear to have activity?

Are there any suspicious people or vehicles in the area?

Since the range of the signal is typically a setting within the administrative function of the router, it would also be possible to lower the signal power, reducing the overall range of the network. This in turn may pull the suspect into closer proximity to the location, making them easier to locate. Recently, I recall an investigation where a neighborhood child was suspect of stealing a laptop from a residence. The network had been secured and would allow access only to that specific laptop, so logging was enabled with e-mail alerts to notify the investigator should any activity be initiated by that laptop. That type of activity alert would notify the investigator that the stolen laptop was in range of the network, which might yield a suspect with evidence in hand.

Wireless Storage Devices

In order to keep up with demand for wireless, many manufacturers now offer remote wireless storage devices, which could pose a significant challenge to investigators trying to locate illegal material. Within the range of a wireless network, a suspect could potentially hide a storage device in an area of their residence that is not readily accessible or apparent. This poses a significant challenge to investigators during consent and search warrant execution. Investigators must always be thinking about the possibility of a remote storage device, especially if it is determined that a wireless network is in use.

Certain limitations with these remote devices can be useful in determining their existence, ultimately helping to determine their location. Even with their portability, these devices will need some type of power source and persistent connectivity to the network. This can limit their proximity to the signal area of the device they typically associate with as well as power availability. When powered, these devices will connect wirelessly to the network they're configured to associate with. This means that if an investigator is able to gain access to the gateway device establishing the network, they might very easily identify that there is another associating wireless device that they have not accounted for. The real challenge comes when these devices are not powered. Without power these devices are off and will not associate with the wireless devices, making them invisible to the entire network. Their discovery would have to come through physical observation at this point, rather then through their virtual presence.

If an investigator had the skills to recognize that a wireless network was in use within the suspect's residence, he or she might be more inclined to ask probing questions about that network, possibly getting the suspect to disclose the existence of a remote device. Physical searches of the residence could also be potentially more productive if the investigator has keyed in on the fact that there might be remote devices involved, requiring a more thorough and educated search.

Interpersonal Communication

As people look to stay connected with friends, family, and coworkers, they are likely to use one or more methods of communication, including e-mail, chat, and blogging—all of which are easily supported on today's computers and portable devices such as laptops, PDAs, and cellular phones. Investigators must be familiar with how these various systems work and how one might be able to retrieve critical case information from stored communications or fragments of previous exchanges. What makes the area of interpersonal communication so important to the investigator is that people are inherently very social; people routinely discuss their daily lives with friends and may even brag about crimes to others. Being able to capture, decipher, and trace back communications to their origin is a critical law enforcement skill.

E-mail

E-mail communication was present at the start of the Internet, and has exploded over the last decade, making it more likely that people today use e-mail in some form or another. E-mail provides another conduit through which people can communicate 24 hours a day, 7 days a week. Unlike a phone conversation that needs the recipient to answer, an active e-mail discussion can be carried out through multiple e-mails spread over time. Messages are sent and are held in a waiting inbox at the convenience of the recipient, who will choose when to read the message and how best to respond. Once an e-mail is read, it is usually up to the receiver to decide and make the conscious choice to delete or discard that communication. This provides a unique opportunity to law enforcement investigating crimes involving e-mails, since undeleted e-mails will be viewable and previously deleted e-mails might be recovered through various forensic methods.

There are countless e-mail addresses and accounts in use today. They fall into two major category types. The first are e-mails generated with e-mail programs that reside on the local user's machine. One of the most common is Outlook or Outlook Express (a Microsoft product), which runs on the user's machine and can be set up with relative ease assuming the account holder has an active Internet connection. E-mails sent and received through this type of account will be stored locally on the user's machine. If this type of e-mail program is used to generate and send illegal communications, it is likely that evidence of those communications might be recovered from the machine used.

The other popular e-mail service is free Internet-based e-mail such as Microsoft's Hotmail and Google's Gmail. These services don't require users to have any special programs in order for them to send and retrieve e-mail in their account. They are able to access mail that is stored on servers provided by the provider they use by signing into a previously created account. These services are extremely portable since they can be accessed from any computer with Internet access and a web browser. With an Internet-based account, an e-mail might be traced back to the originating ISP and it may also be possible to determine the IP address of the machine that connected when the account was created. This is, of course, all dependant upon whether the service provider maintained those records for any specific period of time. Even with this type of account, remnants of Web-based e-mail may be recoverable as HTML documents in temporary Internet files or drive space that hasn't been overwritten by newer files.

With all e-mail cases, it is critical that the investigator follows up on the e-mail address associated with the active case he or she is working. Since there are countless e-mail addresses in use on the Internet, it is not uncommon to have hundreds, if not thousands of variations for the same or similar address. John_Smith@domain is entirely different than JohnSmith@domain. Be sure to match all instances of your suspected e-mail communications exactly.

Chat/Instant Messaging

Chat and instant messaging is another extremely popular method of communication. Unlike e-mail, which ends up being loaded on an e-mail server or downloaded onto the receiver computer's local e-mail program, chats and instant messages are made through direct communication between the two devices. The devices involved exchange communications back and forth in real-time for as long as that "window" is open. Conversations held in chat are not saved by the applications typically used to facilitate this method of communication. This means that for the most part, chat and instant messaging conversations are lost once that session ends.

Service providers do not log chat and instant message traffic, which can be challenging to the investigator investigating a case where chat or instant messaging might have been used. Just like with e-mails, it is extremely important that investigators trace or follow up on the correct screen name or chat id being used by the suspect(s). There are still cases where an investigator might be able to retrieve chat history, as it is possible that one or all of the parties involved may have turn on logging within the application they use. Remnants of chats might also reside on drive space that has not been overwritten by new files. This is where forensic examination can come in very handy if a suspect computer has been seized.

Social Networking and Blogging

Social networking sites, such as Myspace and Facebook, and blogging technologies allow people a conduit through which they can post their thoughts, ideas, and self-expression onto the Internet instantly. For example, within Myspace, users can create an account for themselves along with a personal Web page through which they can express themselves in any manner in which they see fit, be it through music, video, or written expression. These pages become part of a larger online community with similarly minded individuals being able to link together into what is referred to as a *friends network*. Since the information entered at account creation has no true factual verification, it is possible for people to create fictitious identities in order to pass themselves off as someone they're not. The name an investigator might obtain from a Myspace created page might not be the actual identity of the person who created and uses that space. However, it might still be possible to obtain information from the organization responsible for Myspace, such as an account holder's IP address information used during the original account creation or the IP addresses the account holder used to access the account—that type of IP information might be traced back to a suspected user account.

Even though there are no guarantees that information on Myspace pages will be completely factual, this type of online community provides a very powerful and unique service to law enforcement. If an investigator is able to positively identify an online identity as belonging to specific suspect, the investigator might also be able to develop further leads about conspirators based on other identities contained in their friends network. It is critical to investigators that they monitor the activity of potential suspects that they identify by keeping up with the suspect's social networking and blog-related activity.

Media and Storage

Media exists in numerous configurations with varying storage capacities. Most people today are very familiar with the floppy disk, CD-ROM, and DVD—all of which can store and contain files of evidential value. DVDs started reaching capacity sizes in excess of 8 gigabytes, which meant that suspects could save illegal files that would have filled up an entire computer hard drive just years ago on one silver disk. Finding just the right DVD during a search of a suspect or residence could provide numerous evidentiary files. A smaller segment is likely to be familiar with hard drives and understand their role within the computer.

The trend now within media is that of portability. As if trying to find a CD or DVD wasn't hard enough, further technology advances have brought about flash drives and mini smart cards. Many flash drives are smaller than a pack of gum and some mini smart cards are the size of a postage stamp (only thicker) and are capable of holding gigabytes of information. Investigators must be aware of the different types of digital media that exist and be able to identify the media in the field. The variety, and more importantly the size, of media must be taken into consideration when applying for search warrants where digital evidence is suspected; the hiding places for this type of storage are countless.

Summary

What makes computer crime so fearful to some and intriguing to others is the unknown. As investigators learn to deal with and investigate crime involving computers, many are quick to label any crime with a computer presence as a computer/cyber crime. Many of these investigators, and prosecutors, believe that computer crimes are really new crimes; but criminals and "crime" have shown the ability time and time again to be able to adapt to new technologies. It is reasonable to question whether computer crime is just a generational phenomenon caused by a gap in computer understanding and acceptance by many older Americans that did not have the same opportunities to use and learn on computers as the younger generations. Is it likely that this problem will correct itself over time? In the future, computer crime, as it is viewed today, will become nonexistent—not because crime won't exist in the future, but because computer-related crimes will be viewed for what they really are, *crime*.

Frequently Asked Questions

Q: I'm new to cyber crime, but I really want to get involved. Should I jump right into doing forensics?

A: Although there is plethora of training available in the field of digital forensics, you may want to consider getting acclimated to crimes with a cyber component before jumping in with both feet into forensics. Much of what is discussed in this chapter reflects the belief that most cyber crime is just plain ol' crime. Where we may hold this belief to help those that dislike technology realize that they can still work computer crime cases without having a thorough knowledge of computers, we may suggest the same train of thought to you; there is plenty of crime to investigate that has a cyber component that does not require a forensic examination. Tracing e-mail harassments, responding to threats over chat, and investigating sexual solicitations over IM are but a few of the types of crimes that can be investigated without immediately requiring a forensic exam. My recommendation is to find a training course that focuses on the investigation of Internet-related crime—the skills you learn in class such as this won't be wasted if you choose to go the forensics route in the future. By the way, by focusing on crimes that you can investigate without requiring a forensic examination will make your chief a lot happier than your request to purchase $20,000 of software equipment to start processing forensics cases.

Q: I want to get involved with catching predators online. I've seen the TV shows and there doesn't appear to be anything to it. Why should I bother to learn all the technology junk if I don't need to?

A: This is a very popular question. Unfortunately, the fact that it gets asked shows that many people do not know what they do not know, and goes squarely to the heart of *application stupidity*. Agreed, there is little technical knowledge required to "chat" with a potential suspect, and if everything goes according to plan, they show up at your door and you take them into custody. But what happens when things don't go according to plan? Are you aware of the underlying software or process that makes the chatting possible? Is your machine configured correctly and appropriately protected—naming the computer DetectiveDesk22 may show up during a scan of your computer and may blow your cover. Are you knowledgeable about how the particular chatting software works? Does it use a proxy? Will it provide you a direct connection during a file transfer or webcam stream—and if yes, do you have the skills to capture the bad-guy's IP address during that exact moment of transfer? Do you have the skills to properly set up an online identity and protect it from discovery? Although the initial setup of the identity may be trivial, the long-term maintenance and believability of the profile may affect your investigations. In principle, it sounds like a good idea to get a screen name together

to begin enticing predators into the stationhouse, but obtaining basic computer investigative skills will go a long way toward conducting more successful and productive investigations. Further, these skills may prove critical one day when a predator shoots you a webcam image of a child held hostage—that exact moment is not the time to begin learning about the underlying technology—these skills need to be acquired and practiced before employed in active operations.

Chapter 10

Acquiring Data, Duplicating Data, and Recovering Deleted Files

Solutions in this chapter:

- **Recovering Deleted Files and Deleted Partitions**
- **Data Acquisition and Duplication**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

Before data can be analyzed, it needs to be acquired. This means that the data needs to be duplicated so that the person performing the analysis can work from it without modifying the data. This means using any number of tools to duplicate the data so that an exact sector-by-sector mirror image of the disk is generated. This enables the forensic analyst to view any data that is hidden, fragmented, or deleted. Whether data is being duplicated as part of a computer forensic investigation or to acquire a backup of the data for other purposes, it is important that the original data is not modified or corrupted during the duplication process.

Deleted data can be an issue in any situation dealing with computers. A file may be deleted on purpose or by accident, as a normal process of an application, or as the result of a virus, intrusion, or malicious software. In some cases, an entire partition may be lost, causing everything on a volume to appear unrecoverable. However, this is not necessarily the case. When data is deleted, various tools may be used to recover the data from a hard disk or other storage media. In some cases, the files may be corrupted or damaged in some way, and additional software may be needed to repair the file. As discussed in the following sections, regardless of the cause, there are many ways to recover data.

Recovering Deleted Files and Deleted Partitions

At the company where you work, there is probably an inventory of company assets. It may be as simple as a typed list for insurance purposes or as detailed as a database containing records of every desk, computer, printer, and every other asset owned by the company. On your computer, a similar table of records is used to maintain information about the files saved on your hard disk. In the same way that deleting a record from an inventory does not make your desk disappear, deleting a file on your hard disk does not necessarily make the data disappear.

I*Clusters* are two or more sectors of a hard disk, and they are the smallest unit of disk space that can be allocated to store a file. When a file is saved, information on which clusters are used to store the file is kept in another kind of table. A *file allocation table* is used to keep track of files on FAT and FAT32 file systems, whereas on NTFS volumes a *Master File Table* (MFT) is used. Using such tables, the operating system can maintain where files are located. When a file is deleted, the record of the file is removed from the table, thereby making it appear that it does not exist anymore. The clusters used by the file are marked as being free and can now be used to store other data. However, although the record is gone, the data may still reside in the clusters of the hard disk.

Even if multiple partitions are not used, the entire disk can be set as a single partition, formatted to use a particular file system (such as FAT or NTFS), and given a drive letter (such as C: or D:). Information about how partitions are set up on a machine is stored in a *partition table*, which is stored in the Master Boot Record (MBR). When the computer is booted, the partition table allows the computer to understand how the hard disk is organized and then passes this information to the operating system when it is started. When a partition is deleted, the entry in the partition table is removed, making the data inaccessible. However, even though the partition entry has been removed, the data still resides on the hard disk.

As discussed in the following sections, there are many different tools that can be used to recover deleted files and deleted partitions from a hard disk. Depending on how the data was deleted, and whether it was overwritten before it could be recovered, there is the possibility that files and partitions can be restored.

Deleting Files

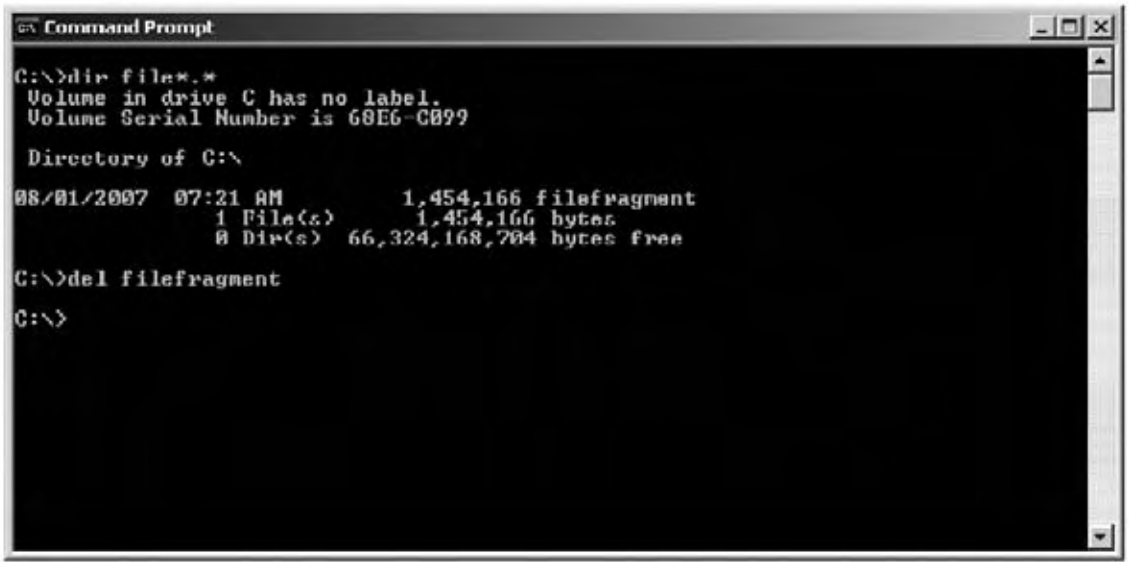
Every operating system (OS) provides a way of removing data from a hard disk. If it did not, a hard disk would fill up quickly from all of the temporary files, backup files, and other data written to a disk by the OS or other software installed on a computer. Although the OS and applications on a machine will generally clean these when they are no longer needed, this is not always the case (as discussed later). In addition to all the files created and deleted by programs on your computer, there are also the ones that you have created and no longer need. These can be old to-do lists, documents that are out of data, and other files that serve no purpose to you anymore. In cases involving computer forensics, it can also be files that contain incriminating evidence.

A *deleted file* is any file that has been logically erased from the file system but may still remain physically on storage media. How a file is deleted can vary. Although for many people, deleting a file means selecting a file and pressing the **DEL** or **Delete** button on their keyboard, there are other ways in which a file may be deleted.

Command Line Delete

In the days before the graphical user interface, the only way to delete a file was from a command prompt. When DOS was one of the major operating systems of choice, you would enter textual commands from a prompt, which was generally a symbol following the drive letter you were currently using. In Windows, the **Command Prompt** is a window that provides an interface to enter the commands to perform certain actions. As shown in Figure 10.1, the **Command Prompt** allows you to enter commands that can be used to delete files and directories from a partition.

Figure 10.1 Command Prompt Delete



There are two commands that can be used to delete files or directories from the command line:

- *DEL*
- *ERASE*

Each of these commands has been available on Microsoft operating systems since DOS and is used to remove a file from any writable storage media. They perform the same function and have the same *switches*, which are options that control what actions are taken by a command. The switches available on systems running Windows 2000, XP, and later are shown in Table 10.1.

Table 10.1 DEL/ERASE Switches

Switch	Description
/p	Prompts you for confirmation as to whether you want to delete the file
/f	Forces the deletion of read-only files
/s	Deletes the specified files from the current directory and any subdirectories
/q	Specifies not to prompt for confirmation before deleting files. This is called *quiet mode*.

Table 10.1 Continued

Switch	Description
/a:	Deletes files based on the following attributes: ■ **r** read-only ■ **a** archive ■ **s** system ■ **h** hidden ■ - prefix used with any of the above attributes, indicating "not"
/?	Displays help

In using this command, you can specify a particular filename or use wildcards. For example, if you wanted to delete a file called filename.txt in the root directory of C:, you would type **DEL C:\filename.txt**. However, if you wanted to delete all files in the current directory, you could type **DEL *.***, which indicates to delete files with any filename or extension.

As previously discussed, the FAT file system uses a file allocation table to keep track of files, whereas NTFS uses an MFT. In using this command, the file that is deleted will have the pointer to that file removed from the table, but the data will remain on the hard disk. This gives you the opportunity to acquire the data using file recovery or forensic tools that are discussed later in this chapter.

Moving Files

Another method of deleting a file involves moving it. If the file is moved from one directory to another on the same partition, the table used to keep track of where files are stored is updated. Because the file still resides on the same partition, a pointer to the file's location is updated. The record showing the file's location is modified to reflect that it is now in another directory, but nothing else about the file changes. Any attributes on the file (such as whether the file is compressed) remain the same.

However, when a file is moved from one hard disk or partition to another, it is actually a multistep process of copying and deleting the file. First, a new copy of the file is created on the target partition. After the file has been copied, the original file is then deleted. This process also requires some housekeeping in the FAT or MFT tables. A new entry is created in the table on the partition where it has been copied, whereas the record for the deleted file is removed from the table on its partition.

When a file is moved from one partition to another, it can offer greater possibilities for recovering a file that has been deleted. If a file is moved to another partition and then deleted later, the file has essentially been deleted twice. As such, you may be able to recover it from the partition on which it was deleted, and you may also be able to recover it from the partition from which it was moved.

Disk Cleanup

Another way in which files are deleted is when software does it for you. Software will generally clean up after itself by deleting old setup, temporary, and backup files when they are no longer needed. However, sometimes files are not properly removed, and a considerable number of files may continue to reside on a hard disk. This is the reason for tools such as Disk Cleanup, which removes files and programs from a Windows XP computer. Disk Cleanup will check for files that can be safely removed so that more disk space is available for use and performance is improved. The Disk Cleanup utility can be run by clicking on the Windows **Start** menu, selecting **Programs | Accessories | System Tools**, and then clicking the **Disk Cleanup** menu item. Once started, the utility will analyze your hard disk to determine what files can be safely deleted, and then it will display a screen similar to that shown in Figure 10.2. As seen in this figure, there are a number of different types of files that may be removed, including the following:

- **Temporary Setup Files**, which are files created by setup programs and no longer used
- **Downloaded Program Files**, which are ActiveX controls and Java programs that were used on certain Web pages you visited
- **Temporary Internet Files**, which are Web pages, images, and other content that are stored on your hard disk so an Internet browser can display them
- **Old *CHKDSK* Files**, which are file fragments that were saved when *CHKDSK* was run to check for disk errors
- **Recycle Bin**, which, as discussed later, is a repository for deleted files
- **Temporary Files**, which are programs temporarily stored in a Temp folder by a program but weren't properly removed by the software. Any temporary files that haven't been modified in the last week are considered safe for deletion.
- **Temporary Offline Files**, which are copies of recently used network files that have been cached, so you can use them after disconnecting from the network
- **Offline Files**, which are local copies of network files that you wanted to have available for use after disconnecting from the network
- **Compress Old Files**, which will compress any files that haven't been recently used
- **Catalog Files for the Content Indexer**, which are files left over from a previous indexing operation, used by the Indexing service

Figure 10.2 Disk Cleanup


Disk Cleanup for (C:)
Disk Cleanup
More Options
You can use Disk Cleanup to free up to 635,632 KB of disk space on (C:).
Files to delete:
Downloaded Program Files 0 KB
Temporary Internet Files 1,745 KB
Offline Webpages 444 KB
Recycle Bin 0 KB
Temporary files 0 KB
Total amount of disk space you gain: 1,745 KB
Description
Downloaded Program Files are ActiveX controls and Java applets downloaded automatically from the Internet when you view certain pages. They are temporarily stored in the Downloaded Program Files folder on your hard disk.
View Files
OK
Cancel

Once you select the types of files you want to delete and press **OK**, the files related to the selected categories are removed from your hard disk. As we saw with previous methods, the entry related to the file is removed from the file allocation table or MFT, and the clusters are marked as free to use for storage of other files.

Permanently Destroying Data

In considering the previously discussed methods, you can see that data can be deleted in a number of ways, but this may not necessarily mean the data is destroyed. If data is deleted, any number of data recovery or computer forensic tools can be used to restore the data. Even formatting the hard disk is not a suitable solution, when you consider that certain tools and data recovery methods can still access the data. The only way to be certain that data cannot be recovered using software solutions is to overwrite it with other data.

Disk erasing software wipes the disk clean by erasing all the files and overwriting the disk space with a series of ones and zeros. In doing so, every sector of the disk is overwritten,

making the data unrecoverable. If anyone attempted to recover data on the disk, they would not be able to retrieve anything because the data is completely destroyed.

A *degausser* or *bulk demagnetizer* is hardware that can be used to destroy data stored on magnetic media such as floppy disks and backup tapes. A degausser is a powerful magnet that erases all data from magnetic media so that no one can retrieve information from it. Hard disks can also have data erased with a degausser, performing a low-level format that erases all data from the disk. If there are concerns regarding particularly sensitive information being seen by outside sources, an additional measure of security is physically scarring or destroying the media. For floppy disks and backup tapes, this involves shredding the media into pieces. For hard disks, you would open the hard drive, remove the platter inside, and physically scar or destroy it. Acid can also be used to destroy magnetic media.

Recycle Bin

So far, we've examined various ways of deleting files by using commands, utilities, and performing actions such as moving files. With each of the methods, the files are no longer visible to the operating system and are considered "permanently" deleted. However, this doesn't mean that the data is completely gone. Until the data is overwritten or destroyed using the methods discussed previously, it may still be recovered. The files are considered permanently deleted by the operating system because there are no programs native to the OS that will allow you to restore the files. For most people, this means the files are simply gone.

To give users a second chance to restore a file before permanently deleting it, a number of operating systems provide an easy-to-use method of restoring deleted files. A *Recycle Bin* or *Trash Can* is a repository where files are temporarily stored after they are deleted. On machines running Windows 95 and later, this container for deleted items is called a Recycle Bin, whereas Macintosh OS uses a Trash Can.

Macintosh OS X allows items to be dragged and dropped into the Trash Can icon on the desktop. By opening the Trash, you can browse previously deleted items, manually drag and drop items out of the trash, or delete all of the items you want to get rid of. Macintosh OS X doesn't provide the ability to delete individual items in the trash; it can only empty everything.

The Recycle Bin has been available in every Microsoft operating system since Windows 95, and it also allows items to be dragged and dropped into the Recycle Bin icon on the desktop. When a file is deleted from the desktop or using Windows Explorer, the file is moved to the Recycle Bin with a record of its original location. The file hasn't actually been deleted; its pointer has simply been changed to show that it now resides in the Recycle Bin. Whereas with Macintosh OS X you must remember where a file was deleted from, Windows will automatically restore an item to its original location.

Another similarity between the two is that you can open the Recycle Bin and browse previously deleted files and directories. Unfortunately, you can't browse the contents of deleted directories in the Recycle Bin. If a file or directory was deleted and you tried to open it in the Recycle Bin to view its contents, double-clicking on it would only display the properties of the deleted item. The option of double-clicking on an item to view its properties isn't available in the Trash Can, nor is the ability to delete individual items from the Trash Can. In the Recycle Bin, however, you can restore or permanently delete individual items so that space can be made available for storing other files.

What Gets Deleted

Although the Recycle Bin is a repository for deleted items, not everything that is deleted will wind up in it. Any items stored on local hard drives that were deleted using Windows Explorer or the desktop will appear in the Recycle Bin. In addition, any items you delete using software that is Windows compliant will send the deleted items to the Recycle Bin. For example, if you were opening a file using Microsoft Word and noticed a file listed in the **Open** dialog box that you wanted to get rid of, selecting the item and pressing the **Delete** key on your keyboard would send the item to the Recycle Bin. However, many programs are not compliant and will simply delete the item, never to be seen again without the use of data recovery or forensic software.

As discussed in previous sections, when you use utilities such as Disk Cleanup, or the *DEL* and *ERASE* commands, the files bypass the Recycle Bin and are simply deleted. The same may occur if software on your machine was cleaning up files after an installation or shutting down. The reason for this is because there are so many files that may be deleted by these programs that it would quickly fill the Recycle Bin. For example, if you used Windows Update to apply a number of fixes or upgrades, the files used to install programs, patches, and service packs would be deleted and never appear in the Recycle Bin. If they did, you would have a difficult time finding the files you mistakenly deleted when browsing the Recycle Bin.

Files are also not deleted if they are not on the local hard drives. If you were to delete a file on a mapped network drive, a compressed folder, a floppy disk, or any other removable storage media, the file would not be sent to the Recycle Bin. The same may occur if a file that is too large for the Recycle Bin is deleted. If there isn't enough space in the Recycle Bin for a file or directory that is deleted, a warning message will appear informing you that it will be deleted. If you agree to the deletion, it bypasses the Recycle Bin and is deleted from the system.

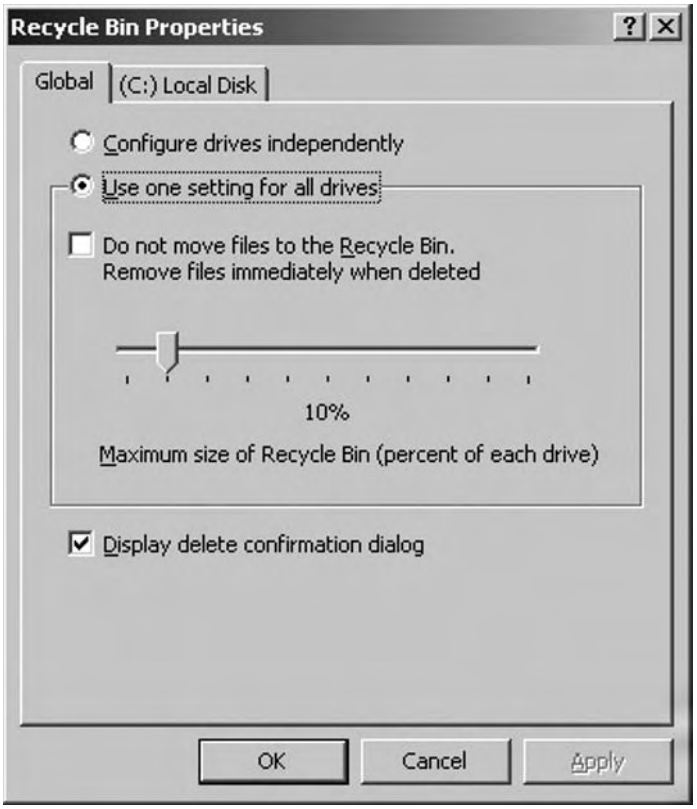
Configuring the Recycle Bin

The Recycle Bin can be configured by right-clicking on the Recycle Bin icon that appears on the Windows desktop and then selecting **Properties** from the context menu that appears. In doing so, a dialog box similar to that shown in Figure 10.3 will

appear. Looking at this figure, you can see there is a **Global** tab and an additional tab for a drive letter. This is because there is one Recycle Bin folder allocated for each partition on a hard disk. In other words, for each hard disk or partition setup on your machine, there is a different Recycle Bin that can be configured and used to store deleted items. By opening the **Properties** of the Recycle Bin, you can either configure all of the Recycle Bins together using the **Global** tab or configure each Recycle Bin separately by using the tab associated with each drive letter. Whether the controls are enabled on each of the drive letter tabs is set using the following options on the **Global** tab:

- **Configure drives independently**
- **Use one setting for all drives**

Figure 10.3 Recycle Bin Properties



The **Global** tab also provides the ability to disable the Recycle Bin so that deleted files are not moved to it but are simply deleted instead. By checking the **Do not move files to the Recycle Bin. Remove files immediately when deleted**, no files that are deleted will be sent to the Recycle Bin.

You can also configure how much space is used to store files in the Recycle Bin using this dialog box. The slider bar on this tab allows you to configure the percentage of disk space that will be used to store deleted files in a Recycle Bin. If you have configured the option of configuring drives independently, then you can set the percentage using a slider that is available on each of the drive letter tabs. On systems prior to Windows Vista, the Recycle Bin is allocated 10% of the disk space for recycled files. This default value can be changed using the slider bar, allowing up to 3.99 GB of space to be used. In Windows Vista, 10% of the disk space can be used on drives up to 40 GB, but above that, Windows Vista allows a maximum of 4 GB plus 5% of the disk capacity. If the amount of data stored in the Recycle Bin exceeds this amount, then the oldest files are removed (permanently deleted) from the Recycle Bin first so that newly deleted files can be stored there. This is referred to as FIFO, or *first in, first out*.

Storage Locations of the Recycle Bin

When files are deleted, they are stored in a special hidden directory, which is located in the root directory. Because one Recycle Bin is allocated for each partition on a computer, there is one of these folders on each partition of a hard disk. The name of this folder depends on the file system and operating system being used:

- On FAT file systems, the folder is named *Recycled*. For example, on your C: the storage location of the Recycle Bin would be C:\RECYCLED.
- On NTFS file systems used by Windows NT, 2000, and XP, this folder is named *Recycler*. For example, on your C: the storage location of the Recycle Bin would be C:\RECYCLER.
- On NTFS file systems used by Windows Vista, this folder is named *$Recycle.Bin*. For example, on your C: the storage location of the Recycle Bin would be C:\RECYCLE.BIN.

If you were to search these directories on a partition of your hard disk, you might be surprised to find that none of these appear. The reason is probably because they are hidden folders. To view these directories, you would need to configure the View settings in Windows Explorer. In Windows Explorer, click on the **Tools** menu, and then click on the **Folder Options** menu item. When the **Folder Options** dialog box appears, click on the **View** tab, and uncheck the **Hide protected operating system files (Recommended)** checkbox. Once this box is unchecked, you should be able to view the hidden folders through Windows Explorer or My Computer.

On NTFS partitions, a Recycle Bin is created in the Recycler folder for each user who logs onto the computer. Because the Recycler directory contains Recycle Bins for each user of the computer, a number of subdirectories will exist under this folder. Each of the subdirectories is named after the Security ID (SID), which is a unique number that identifies users and groups in Windows, appearing similar to S-1-5-21-191058668-193157475-1542849698-500. Although opening the Recycle Bin will automatically show the contents of a directory related to the account you logged in with, looking at the subdirectories in the Recycler folder will require you to find which SID corresponds to your account.

NOTE

For a listing of well-known security identifiers in Windows, you can visit support.microsoft.com/kb/243330. This will provide you with information on SIDs associated with particular groups and users. For example, if you were analyzing the Administrator account, this would be a well-known SID and could be identified as S-1-5-*domain*-500.

Undeleting or Permanently Deleting a File

Before we discuss how to undelete or remove a file from the Recycle Bin, let's review exactly what happens when the file is deleted. When you delete a file from the desktop or using Windows Explorer, My Computer, or a compliant application, the entry in the file allocation table or MFT is removed, and the file is indicated as being located in the Recycle Bin directory in the root drive of the partition on which it was deleted. The file is then renamed using the following syntax:

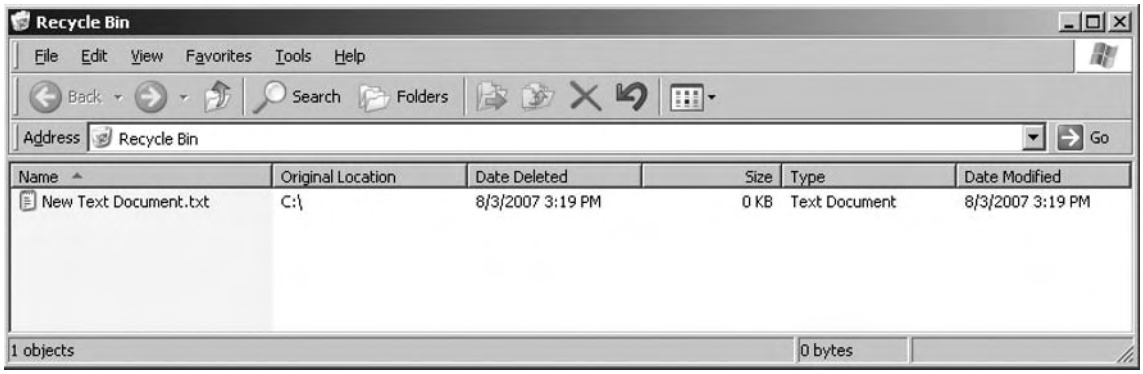
D<*original drive letter*><#>.<*original extension*>

To illustrate how the renaming works, let's say you deleted C:myfile.txt, and this was the second file that was deleted and residing in the Recycle Bin. In using the previous syntax, the file would be renamed Dc2.txt. By examining this, you can see that it starts with the letter D, followed by the original drive letter (c), that it was the second file deleted, and that it uses a .txt extension. Its original name and location are stored in a hidden index file named INFO, which is located in the Recycle Bin folder in the root directory. On Windows 95 machines and partitions using NTFS, this file is called INFO2.

The information stored in the INFO file is important because it allows the file to be automatically restored to its original location. You can view information on the files that have been deleted and stored in the Recycle Bin by double-clicking on the Recycle Bin

icon on the Windows desktop. In doing so, a screen similar to that shown in Figure 10.4 will appear. This figure shows that the Recycle Bin allows you to view information such as when a file was deleted and its original location. By right-clicking on a file, you can then click on the **Restore** menu item in the context menu that appears. By restoring a file, the original path can be read from the INFO file, and the file can be renamed and restored to its original path and filename.

Figure 10.4 Recycle Bin



Deleting or emptying files from the Recycle Bin permanently deletes them from the system, so they can be recovered only with data recovery or forensic software. To delete a file from the Recycle Bin, you can select the file and then press the **DEL** or **Delete** key on your keyboard. You could also delete all of the files from the Recycle Bin by right-clicking on the desktop icon and then clicking **Empty Recycle Bin** from the context menu that appears.

Although the Recycle Bin can be configured so that files are immediately deleted, you can also delete a file and not have it go to the Recycle Bin on a file-by-file basis. If you wanted to bypass the Recycle Bin and simply delete the file, you would select the file you want to delete and then hold down **Shift +DEL** (that is, the **Shift** key and **Delete** key at the same time). When doing this, a prompt will appear asking you to confirm the deletion, and upon choosing **Yes**, the Recycle Bin is bypassed and the file is deleted.

Damaged Recycle Bins

As with any other software, there are times when the Recycle Bin might not function correctly and needs to be repaired. If you have deleted files and the Recycle Bin appears to be empty, it could mean that something is wrong with the INFO file. As we discussed, the INFO file maintains information on files stored in the Recycle Bin. If this is damaged or deleted, no files will appear in the Recycle Bin, but this doesn't mean that anything has happened to the deleted files. Although the files were renamed when they were put in the

Recycle Bin folder, they contain the same data and will appear under the Recycled folder. To restore the file, you could simply rename the file to its original name.

Repairing a deleted INFO file is one of the easiest repairs you'll encounter because it will be re-created when you restart Windows. If it is damaged, however, you will need to delete the INFO file from the Recycled folder. Once deleted, you would then restart Windows so it can be re-created. To delete an INFO file, you would follow these steps:

1. From the Windows **Start** menu, select **Programs | Accessories** and then click on the **Command Prompt** menu item.
2. Type **cd recycled** if it is a FAT partition or **cd recycler** if it is NTFS. On an NTFS partition, you will also need to navigate into the directory named after the SID related to the user account you logged in with. Press the **Enter** key.
3. Type **attrib –h inf*** and the press the **Enter** key.
4. Type **del info** (or **del info2** on NTFS partitions) and then press the **Enter** key.
5. Restart your computer.

Another problem that may occur with the Recycle Bin is that the RecycledRecycler directory may be damaged. When the folder is damaged, no files appear in the Recycle Bin. When this occurs, you would create a copy of the DESKTOP.INI file, which is located in the Recycled directory or in the Recycler subdirectory that matches the SID of the user who is currently logged on. After copying this file, you would then delete the contents of the Recycled directory or the appropriate subdirectory in the Recycler directory. Doing so deletes all files in the Recycle Bin. Once this is done, you would then copy the DESKTOP.INI file back to its previous location.

The final part of the Recycle Bin that may possibly be damaged is the Recycled\Recycler folder. When this happens, the Recycle Bin icon on the desktop will appear full, but the contents cannot be viewed and the **Empty Recycle Bin** command (which appears on the context menu when you right-click on the Recycle Bin icon) will be disabled. To fix this problem, you would perform the following steps:

1. From the Windows **Start** menu, select **Programs | Accessories** and then click on the **Command Prompt** menu item.
2. Type **attrib –s –h recycled** if it is a FAT partition or **attrib –s –h recycler** if it is NTFS. Press the **Enter** key.
3. Type **del recycled** if it is a FAT partition or **del recycler** if it is NTFS. Press the **Enter** key.
4. When prompted as to whether you are sure, press the **Y** key.
5. Restart your computer.

Data Recovery in Linux

Just as files can be removed in Windows and Macintosh operating systems by selecting a file and pressing the **Delete** key on the keyboard, Linux files can also be deleted from a hard disk in this way. Files can also be deleted in Linux and UNIX machines by using commands. To delete a file in Linux or UNIX, you would use the *rm* command. Deleting a file with this command won't prompt you for any confirmation by default, so once it is deleted, you will have to rely on data recovery software to restore the file. Like the *DEL* and *ERASE* commands discussed previously, there are a number of switches that can be used with the *rm* command (Table 10.2).

Table 10.2 *rm* Switches

Switch	Description
-f	Forces deletion of files, and ignores nonexistent files, without prompting
-i	Interactive mode in which confirmation is required before files are deleted
-r	Contents of a directory removed recursively
-v	Verbose; provides details of what is done

A more secure way of deleting files in Linux and UNIX is to use the *shred* command. With this command, the file is deleted and overwritten to ensure that it can't be recovered. As mentioned previously, by overwriting the data on a disk, the file cannot be recovered. As such, it is important that this command is only used when you are absolutely sure that you want the file to be destroyed. There are a number of options that can be used with this command, as shown in Table 10.3.

Table 10.3 *shred* Switches

Switch	Description
-f	Forces deletion of files and will change permissions to allow writing if necessary
-n	Iterations in which the file will be overwritten. By default, it is overwritten 25 times, but with this command, you can specify a number.
-s	Shreds a specific number of bytes
-u	Truncates and removes the file after writing
-v	Verbose; provides details of what is done
-x	Indicates not to round file sizes to the nearest block
-z	Adds a final overwrite with all zeros to hide that the file was shredded

There are also a number of desktops that can be used on Linux that provide a Recycle Bin that can be used to restore files that were deleted by selecting a file and pressing the **DEL** key on your keyboard. As mentioned previously, Recycle Bins are part of the desktop environment, and on Linux and other operating systems, a number of desktop environments can be used, such as GNOME and Xfce. These desktop environments provide a Recycle Bin, allowing files to be restored or permanently deleted. However, any files deleted using the *rm* or *shred* command bypass the Recycle Bin. Whereas the *shred* command overwrites the data so it cannot be restored, the *rm* command does not modify the data after deleting it. This means that any files deleted using the *rm* command could have data restored using recovery tools.

Recovering Deleted Files

When a file is deleted, it doesn't necessarily mean that the data cannot be completely or partially recovered. Data written on a hard disk generally stays there unless or until it is either overwritten by more data or physically erased by a magnet. Simply deleting the data using operating system file management utilities does not get rid of the data. It only removes the pointer used by the file system to locate that data physically on the disk. The data itself (in the form of the physical changes to the disk's magnetic surface) is still there and can be recovered using special recovery software.

Many people also think that formatting a hard disk erases all its data, but this isn't necessarily so. *Formatting* defines the structure of the disk. *Low-level formatting* (LLF), which physically defines where the tracks and sectors are on the disk, does erase data. However, modern disks are formatted at the low level at the factory; users do not perform LLF on today's IDE and SCSI disks. Thus, when we discuss formatting, we are generally talking about *high-level formatting*. This term refers to the process of defining the file system structure. Thus, we say a disk is formatted in FAT or formatted in NTFS.

Assuming a file has not been overwritten or damaged in some way, there is always the chance that tools can be used to restore the data. The question that will determine which tools you use is, *Why am I recovering it?* This question in itself will decide whether you will need to perform data recovery or computer forensics.

Although the two terms are often used interchangeably, there is a difference. *Data recovery* is a process of salvaging data that was lost or deleted. *Computer forensics* is a process of gathering and examining evidence to establish facts so that accurate testimony and evidence can later be presented in court or other hearings. In considering these definitions, you will notice that there is a difference in why the data is being acquired: Data recovery seeks to restore the data, whereas computer forensics seeks to obtain data to be used as evidence in court.

Mistakes happen, and its common for a person to accidentally delete a file needed for work or personal use. Ask any network administrator, and they'll tell you that there are many occasions when they need to restore a file that a user has deleted. On personal computers,

workstations, and laptops, you can restore the data from the Recycle Bin or Trash Can. If the file has been deleted from removable media or has been emptied from the Recycle Bin, then data recovery software can be used to retrieve the data.

Such software is not suitable in other situations. Standard data recovery software may not guarantee that the file won't be modified when it is recovered, or the software may generate temporary files that could damage other data on the disk. To preserve data, computer forensic software is necessary.

If you are investigating an incident, always assume that the case will eventually go to court. If you used software that was not designed for forensic use, you could alter or damage data on the disk and compromise the investigation. In addition, any data that you retrieved could be deemed inadmissible in court. In any situation in which you are investigating an intrusion, policy violations, cybercrimes, or other incidents in which data could be used as evidence, you should always take the side of caution and use computer forensic tools.

Regardless of which type of tool you use, you should never install the software on the drive containing the data you wish to recover. For example, if you accidentally deleted a file on your C: and then downloaded and installed data recovery software to your C:, you could overwrite the data that you were trying to recover in the first place. Computer forensic software may use bootable disks to access the drive or connect to the machine using a network cable. By sharing the drive and connecting to it over a network or using a network cable to connect two machines, you could restore the data without worrying about corrupting data on the disk. If the computer has two drives and you are performing data recovery, you could also install the software on another drive. However, if you deleted the file on a boot partition, files may be installed in your operating system's directory or other directories used to store common files. When using these tools, remember that the integrity of the data you are seeking to recover is paramount, and take whatever actions necessary to keep it from being damaged or destroyed.

Deleted File Recovery Tools

Data recovery tools are designed to restore data that has been deleted or corrupted from any number of sources, including hard disks, CDs, DVDs, Blu-ray, HD-DVD, floppy disks, memory cards used in digital cameras, and other storage media. Depending on the capabilities of the software, it will scan the media and search for any damaged, corrupted, or deleted files and display which ones are available for recovery, allowing you to choose which ones will be restored. In some cases, the tools will even repair damaged files so that data can be accessible again.

Undelete Tools

As mentioned previously, using commands such as *DEL* and *ERASE* from the command line or holding down the **Shift** key when deleting a file will bypass the Recycle Bin. To restore the file, you need to use tools that will search the hard disk for deleted files and allow

you to undelete them. Fortunately, there are a number of tools available that perform this task, with various features that make undeleting files easier. They should be used as soon as possible, however, to avoid any other data overwriting the file.

Undelete

In discussing tools that undelete files, we start with one of the most basic ones used in MS-DOS 5.00 to 6.22. *Undelete* is a command-line tool that can be used on these old systems, allowing you to simply type **UNDELETE** followed by the path to the file you wish to restore. For example, if you wanted to undelete a file named mytext.txt from your root directory, you would simply type **UNDELETE C:\MYTEXT.TXT**. Doing so would restore the file. In addition, there are a number of switches that can be used (Table 10.4).

Table 10.4 UNDELETE Switches

Switch	Description
/all	Automatically restores all of the specified files
/list	Lists all files that are available to be restored
/DOS	Limits recovery to files that meet the file specifications in the disk directory table
/DR	Limits recovery to files in the Delete Tracking File
/DS	Limits recovery to files in the SENTRY directory
/load	Loads UNDELETE.EXE into memory using information in the UNDELETE.INI
/unload	Unloads UNDELETE from memory
/purge	Deletes contents of the SENTRY directory
/status	Displays the type of delete protection on the computer
/S	Loads the memory resident portion of UNDELETE and enables the Delete Sentry level of protection
/T	Loads the memory resident portion of UNDELETE and enables the Delete Sentry level of protection. Optional entries may be used to specify the maximum number of entries found in the PCTRACKR.DEL file.

People who have been using computers since the days of MS-DOS tend to automatically think of this program when mentioning *Undelete*. However, it was designed to work only on systems running MS-DOS 5 to 6.22, and it can cause problems when run on systems running Windows 9*x* or later. There are, however, a number of other programs that use this name but are designed for newer operating systems.

Active@ Data Recovery Software

Active@ Data Recovery Software is a company that has developed a number of tools used to recover lost or deleted files from storage media. Two of these are

- Active@ UNDELETE
- Active@ UNERASER

Active@ UNDELETE is data recovery software available from www.active-undelete.com. Using this tool, you can recover data from basic and dynamic volumes, including RAID volumes, and large hard disks that are more than 500 GB in size. It also supports recovery from removable storage media such as USB flash drives, ZIP drives, and memory sticks and cards.

Active@ UNERASER is another data recovery product available from www.uneraser.com, and it will run on Windows or DOS. Using this tool, you can access deleted files before Windows even starts, and it supports local files, compressed files, MBR backups and can access sectors of the disk drive with a disk viewer feature. The Active@ UNERASER tool can run from either a bootable floppy disk or a CD. It also provides a Bootable ISO CD Image and Bootable Floppy Creator to create the disk or CD that can then be used to recover files.

R-Undelete

R-Undelete from r-Tools Technology (www.r-undelete.com) is a tool that restores deleted files, but it also provides an easy-to-use wizard that takes you through the steps of recovering a file. In addition, it provides features that allow you to reconstruct damaged graphics, audio, and video files. Before recovering a file, you can preview the file to determine if you actually want to restore it or leave it deleted.

Easy-Undelete

Easy-Undelete (www.easy-undelete.com) is a tool that will not only restore files from hard disks using FAT12, FAT16, FAT32, and NTFS file systems but also restore files from non-Microsoft partitions such as Linux and Macintosh OS X. It also supports other storage media, such as memory cards used in digital cameras, including a preview feature that allows you to view images before restoring them. In addition, it provides a hexadecimal preview feature that allows you to view the contents of clusters.

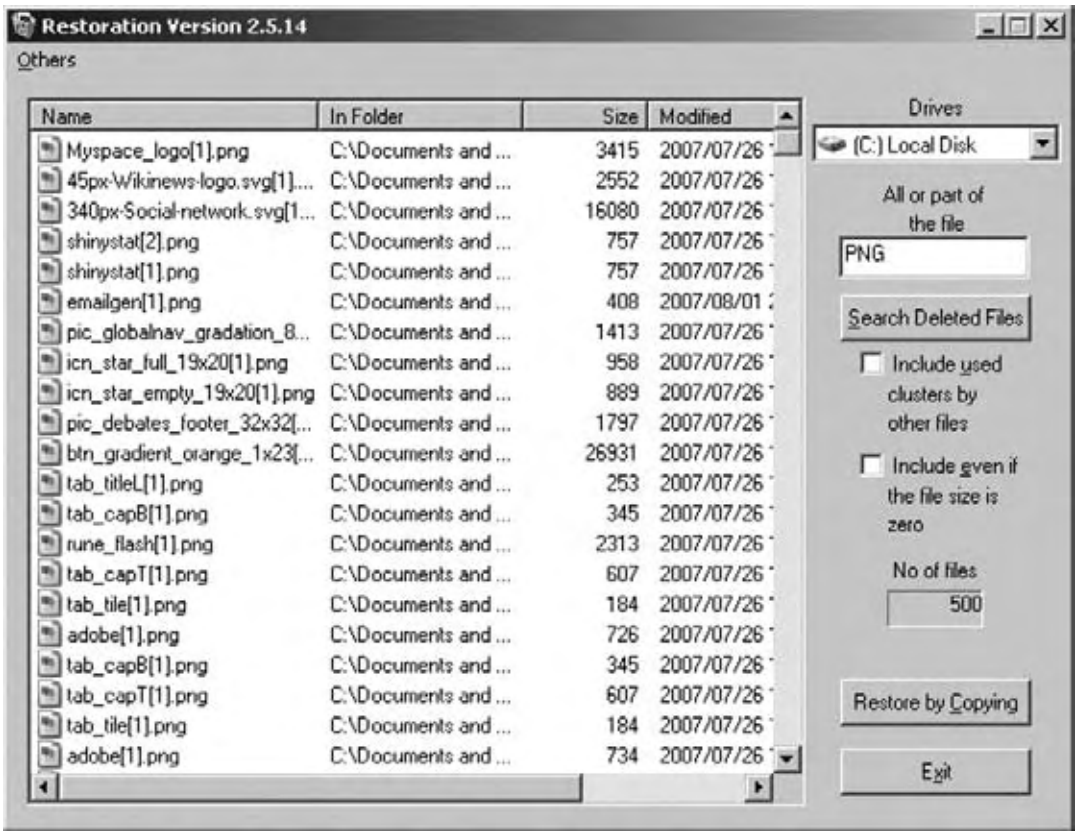
WinUndelete

WinUndelete is a popular recovery tool that is available from www.winundelete.com. It allows you to recover files from Microsoft file systems with the original create and modified storage dates. It provides a search feature to scan for specific files, and it allows you to filter results by extensions and file types. It also allows you to preview certain types of data before restoring them, such as Microsoft Office documents, images, and plain text.

Restoration

Restoration is a free tool that was written by Brian Kato and is available for download from software distribution sites such as www.snapfiles.com. It is a fast and easy-to-use tool, which is small enough to run from a floppy disk. As shown in Figure 10.5, it not only allows you to view any deleted files on a hard disk but also allows you to search for specific files by entering all or part of the filename into a search box.

Figure 10.5 Restoration


Restoration Version 2.5.14
Others
Name
In Folder
Size
Modified
Myspace_logo[1].png C:\Documents and ... 3415 2007/07/26
45px-Wikinews-logo.svg[1].... C:\Documents and ... 2552 2007/07/26
340px-Social-network.svg[1... C:\Documents and ... 16080 2007/07/26
shinystat[2].png C:\Documents and ... 757 2007/07/26
shinystat[1].png C:\Documents and ... 757 2007/07/26
emailgen[1].png C:\Documents and ... 408 2007/08/01
pic_globalnav_gradation_8... C:\Documents and ... 1413 2007/07/26
icn_star_full_19x20[1].png C:\Documents and ... 958 2007/07/26
icn_star_empty_19x20[1].png C:\Documents and ... 889 2007/07/26
pic_debates_footer_32x32[... C:\Documents and ... 1797 2007/07/26
btn_gradient_orange_1x23[... C:\Documents and ... 26931 2007/07/26
tab_titleL[1].png C:\Documents and ... 253 2007/07/26
tab_capB[1].png C:\Documents and ... 345 2007/07/26
rune_flash[1].png C:\Documents and ... 2313 2007/07/26
tab_capT[1].png C:\Documents and ... 607 2007/07/26
tab_tile[1].png C:\Documents and ... 184 2007/07/26
adobe[1].png C:\Documents and ... 726 2007/07/26
tab_capB[1].png C:\Documents and ... 345 2007/07/26
tab_capT[1].png C:\Documents and ... 607 2007/07/26
tab_tile[1].png C:\Documents and ... 184 2007/07/26
adobe[1].png C:\Documents and ... 734 2007/07/26
Drives
(C:) Local Disk
All or part of the file
PNG
Search Deleted Files
Include used clusters by other files
Include even if the file size is zero
No of files
500
Restore by Copying
Exit

Mycroft V3

Mycroft V3 is computer forensic software developed by Dibs USA and is available at www.dibsusa.com. The software runs from a bootable floppy disk, and it provides a search engine that is used to scan a computer for data on the disk. Using this tool, an investigator can determine if there is evidence on the computer. By using this tool to provide a cursory search of a computer's hard disk, you can determine if additional investigation and acquisition of evidence is necessary.

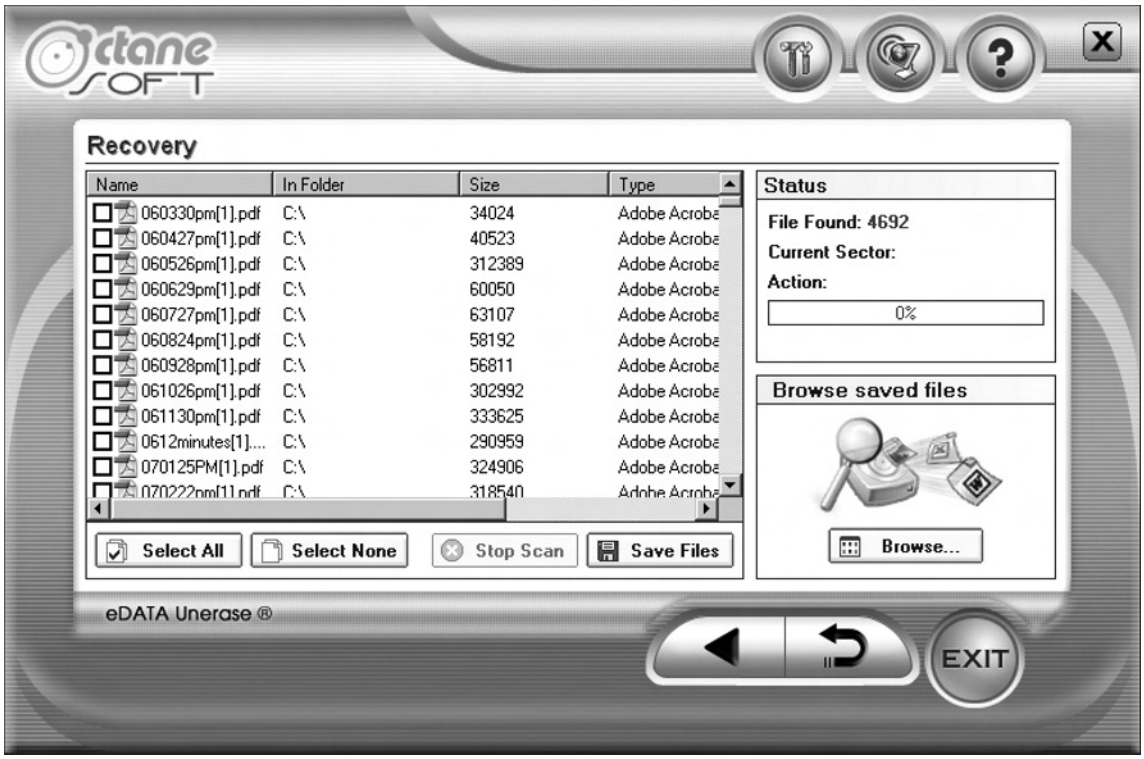
Recover My Files

Recover My Files is a tool created by GetData Software Development (www.getdata.com) and is available for download at www.recovermyfiles.com. It provides an easy-to-use interface that allows you to recover data from hard disks and removable storage media such as floppy disks, Zip drives, USB, and camera cards. As discussed later, this is just one of the tools available from GetData that can be used for data recovery.

eData Unerase

eData Unerase is a free tool developed by Octanesoft and is available at www.octanesoft.com. It allows you to restore any files that have been deleted from the Recycle Bin or using other methods that make it otherwise unrecoverable without data recovery software. It provides an easy-to-use interface with a wizard that takes you through the steps of recovering your data, and it allows you to specify whether you want to search for all deleted files or just specific types of files (for example, pictures, multimedia, documents, and Web pages). If you know the filename, you can also enter all or part of the name to search for. As shown in Figure 10.6, once the wizard is complete, you simply select which files you'd like to restore and then click the **Save Files** button to specify where the recovered data should be saved.

Figure 10.6 eData Unerase



Recover4all Professional

Recover4all Professional is a tool for recovering data from FAT12, FAT16, FAT32, NTFS, and RAID volumes, and it is available at www.recover4all.com. It also allows recovery from encrypted drives that have not been encrypted with EFS, compressed drives, and Windows ME/XP protected files. Its interface is easy to use, appearing similar to Windows Explorer; folders appear in the left pane, and selecting a folder will allow you to view deleted files in the right pane. By selecting a file and clicking the **Recover** button, you can specify where on the disk you want to restore the file.

File Scavenger

File Scavenger is a tool developed by QueTek Consulting Corporation and is available at www.quetek.com. It allows data to be recovered from hard disks formatted as FAT12, FAT16, FAT32, and NTFS. It can also recover data from floppy disks, ZIP disks, memory sticks, flash cards, RAID volumes, compressed drives, and other media. It provides data recovery features using an interface similar to Windows explorer, and it allows you to search for specific files using search strings and wildcards.

VirtualLab

VirtualLab is data recovery software designed for use on Windows 9*x* or later, as well as Apple Macintosh computers running OS 9 or later. It is available at www.binarybiz.com, and it allows you to recover files from drives using FAT12, FAT16, FAT32, NTFS, HFS, or HFS Plus partitions. It can also be used to recover data from other media, such as Apple iPods, digital camera cards, and removable storage. Using an exhaustive scan, it can locate lost files and will also locate lost partitions. Using this tool, you can not only undelete files but also unformat partitions to restore data from a formatted, damaged, or lost partition. Once the files you want to retrieve are found, they can be saved to another location (such as a network drive or another local disk).

File Recover

File Recover is a tool developed by PC Tools Software that is available at www.pctools.com/file-recover/. It supports FAT16, FAT32, and NTFS file systems and also IDE/ATA/SCSI hard drives, including drives larger than 8 GB. It provides an easy-to-use interface that allows you to quickly scan for files and also advanced features that allow you to search by all or part of the filename, specific extensions, or file size and scan either the logical drive or the physical hard disk.

As shown in Figure 10.7, once a scan is complete, it will provide a listing of deleted files, with information on the condition of each file that was found. Double-clicking on an image file will provide a preview of the image, allowing you to view the picture before restoring it. The left pane of the interface also allows you to view different files that were found based on file type. Clicking on items in the left pane will display only HTML documents, PNG images, JPEGs, or files in other file formats.

Figure 10.7 File Recover



Badcopy Pro

Badcopy Pro is another popular data recovery tool that was developed by Jufsoft and is available at www.jufsoft.com/badcopy/. It can restore data from hard disks, damaged or defective CDs and DVDs, floppy disks, USB flash drives, as well as Iomega Zip, Jas, and MO disks and other storage media. It has features to repair damaged floppy disks, recover photos from memory cards used in digital cameras, and data recovery on UDF and packet-writing disks. One of the benefits of this tool is that it doesn't write data to the disk or media where data is being recovered, and it saves recovered data to a location that you specify. This prevents overwriting any data on the disks, which could prevent further recovery.

Zero Assumption Recovery

Zero Assumption Recovery is a recovery tool available at www.z-a-recovery.com. It provides support for FAT16, FAT32, and NTFS, RAID0 and RAID5 hardware recovery, as well as limited support for the Linux ext2 file system. To perform recovery, it uses a wizard-style interface that will take you through the process of determining how and where data is to be recovered. Recovery can take a considerable amount of time, as it can scan the physical disk,

locating bad sectors and data fragments in addition to complete files that have been deleted. As discussed later, it also provides features that allow you to recover images from memory cards used by digital cameras.

SUPERFileRecover

SUPERFileRecover is another easy-to-use tool to recover files, which is available at www.superfilerecover.com. It allows you to search for all files on a partition, or specific filenames, supporting wildcards. You can also specify to search only for music, movies and video, pictures and photos, or documents. Once a deleted file is selected in the interface, you can view the file in a preview pane.

DiskInternals Uneraser and NTFS Recovery

DiskInternals (www.diskinternals.com) provides a number of data recovery solutions, each of which provides an easy-to-use intuitive interface that provides wizards for restoring data. Two of the tools available from DiskInternals are

- NTFS Recovery, which restores data on NTFS drives that are corrupted, formatted, or otherwise inaccessible. It will also recover data if the partition table has been corrupted or damaged. It provides a wizard that will scan the disk and then restore the original structure of folders and files.
- Uneraser, which is designed to recover files that have been deleted from FAT/FAT32 and NTFS drives, and it will also recover encrypted files. It supports recovery from compressed drives, USB flash drives, digital cameras, as well as other storage media.

PC Inspector File Inspector

PC Inspector File Recovery is a free tool developed by CONVAR that can be used to restore files from partitions formatted in FAT12, FAT16, FAT32, and NTFS. It will restore files with their original timestamps, and it allows recovered files to be saved on a network. It is available at www.pcinspector.de/file_recovery/uk/welcome.htm.

Search and Recover

Search and Recover is a data recovery tool from Iolo Technologies (www.iolo.com) that is designed to restore documents, images, music files, and other data from a variety of storage media. The tool provides wizards that take you step by step through the recovery process, and it allows you to preview documents and multimedia before restoring them. To prevent any data loss from installing the software, Search and Recover can run from the installation CD, and it provides features to make a disk image that will duplicate any data on the hard disk, CDs, or DVDs.

Additional features in the software also allow you to permanently delete the data by having files overwritten by other data. You can configure Search and Recover to overwrite the files up to 100 times, making these files unrecoverable with any data recovery software. It also provides features to recover deleted e-mails from Netscape, Thunderbird, Eudora, Microsoft Outlook Express, and Microsoft Outlook. In addition, you can restore deleted e-mail, tasks, notes, and other elements in Outlook.

O&O Unerase

O&O Unerase is a tool developed by O&O Software to recover data from hard disks, floppy disks, digital cameras, memory cards, and other media. It provides an easy-to-use interface that simplifies finding and recovering files. Using the search options in this tool, you can search for deleted files by name, type, size, or date of last change. It is available at www.oo-software.com/en/products/oounerase/.

Filesaver

Filesaver, available at www.file-saver.com, is a tool that can recover data from FAT and NTFS drives, floppy disks, memory cards used with digital cameras, and other storage media. It provides different modes, including one that allows you to view deleted files through a Windows Explorer-style interface, and allows you to preview images before restoring them. Using this tool, you can not only restore individual files but also restore entire folders of data.

Stellar Phoenix

Stellar Phoenix is a data recovery tool from Stellar Information Systems and is available at www.stellarinfo.com. It supports FAT and NTFS file systems, and it will analyze the hard disk for damage and corrupted data before offering you a listing of files available to restore. It will also allow you to recover data from hard disks that have been formatted, even after an operating system has been installed, and provides features for recovering data from repartitioned and lost hard drives.

Restorer 2000

Restorer 2000 is a data recovery tool available at www.restorer2000.com that supports recovery from hard disks using FAT12, FAT16, FAT32, and NTFS file systems, as well as floppy disks, memory cards used in digital cameras, and other storage media. It will recover data from IDE, ATA, SATA, and SCSI hard drives and will recover and reconstruct data from RAID5, damaged volume sets, stripe sets, and mirrored volumes. It will also recover data from damaged partitions and hard disks that have been formatted, and it will search for deleted files by name, date, mask, size range, attributes, and other search parameters. If a graphic, audio, or video file is being recovered, you can preview the file before restoring

it and use a built-in hexadecimal editor to view and edit files. If the network version of this software is used, then you can also connect to remote computers, scan their hard disks and partitions, and restore data across the network.

R-linux

R-linux is a free data recovery tool developed by r-Tools Technology and is available at www.data-recovery-software.net/Linux_Recovery.shtml. Unlike the other tools discussed in this section, R-linux is specifically designed for the ext2 file system used on Linux and UNIX machines. It can recover data from local disks, but it is unable to reconstruct damaged RAID or stripe sets. In acquiring data from a disk, it can create an image file of the entire disk or partition and allow data to be acquired from the image file as if it were a regular disk. Once a file is recovered, it can be saved to a local or network drive that you specify.

PC ParaChute

PC ParaChute is a backup tool developed by UniTrends and is available at www.unitrends.com. It is used to perform a backup of the first physical disk and partitions on the computer. In doing this, it writes your system's boot drive and creates an image of the computer on a server backup tape, ensuring that any data that is lost can be recovered by restoring the backed up image. Restoring the data involves booting from a recovery disk and then choosing a restore option from a menu. PC ParaChute automatically the image file and restores the machine.

Although most of the tools in this section deal with restoring deleted files from data on the disk, it is important to realize that backing up data is the easiest and most reliable method of restoring data. By making a copy of any data on a hard disk or other media, you can then restore it as needed. Regular backups will ensure that recent versions of files are ready as needed so that if there is a hard disk failure or some other problem, the data can be quickly restored to its previous state.

Recycle Bin Replacements

An ounce of prevention is worth a pound of cure, and if users have a way of recovering their own files, it means you won't have to do so. Although the Recycle Bin is a wonderful tool to give users a second chance before permanently deleting a file, it isn't without its flaws. As previously mentioned, not all files are sent to the Recycle Bin, so some files aren't recoverable by the common Windows user. Because developers found ways to improve the security of the Recycle Bin, and allow better opportunities to restore files, there are a number of products that can be installed on Windows machines to replace the existing Recycle Bin.

Diskeeper Undelete

Undelete from Diskeeper Corporation (www.undelete.com) is a tool that runs on Windows XP and 2000, and it replaces the Recycle Bin with a *Recovery Bin*. Once installed, files that are deleted are sent to the Recovery Bin, allowing you to search the bin and restore any files

that were accidentally deleted. It even provides the ability to view versions of files that were deleted in order to select the correct version to restore. If you had a Microsoft Word, Excel, or PowerPoint file that was overwritten, you could use the Recovery Bin to restore a previous version. Versions of the Undelete software can be used to protect local files, whereas the server edition can be used to protect and restore files deleted by network clients.

Fundelete

Fundelete is a tool that replaces the Recycle Bin on systems running Windows NT, ME, 2000, and XP so that any files that are deleted from the **Command Prompt**, using the **Shift + DEL** command, or within a program can be recovered. It also provides filter options so that files with specific file extensions are not sent to the Fundelete Bin.

Fundelete was created by Sysinternals, which was acquired by Microsoft in 1996. As such, visiting their Web site will redirect you to Microsoft's Web site. Information on the Sysinternals products can be fount on Microsoft's TechNet Web site at www.microsoft.com/technet/sysinternals/default.mspx, whereas copies of Fundelete are still available from sites such as http://fundelete.en.softonic.com.

CD/DVD Data Recovery

As you are well aware at this point in the book, not all data is stored on hard disks, so it follows that there would be data recovery software that is specifically designed to restore damaged and deleted files stored on CDs and DVDs. When data is stored on CDs and DVDs, the data can not only be deleted from rewritable discs but also be damaged by scratches, other damage, and defects in the media. The following are some of the more popular programs used to restore data:

- CDRoller
- IsoBuster
- CD Data Rescue
- InDisk Recovery

CDRoller

Despite its name, CDRoller recovers data written to both CDs and DVDs and is available at www.cdroller.com. It not only allows you to recover data from CD-ROM, CD-R, and a variety of other CD and DVD formats but also allows you to recover data recorded directly onto the discs, such as that written by stand-alone DVD recorders, camcorders, and other products. It provides the ability to split recovered VOB and VRO files into separate clips, and it can convert raw video into MPEG files, allowing you to create a new video from damaged data.

This tool also provides software to burn data onto CDs and DVDs, allowing you to back up any data before there is a problem or once its been recovered. It supports the ISO 9660 file system, Joliet extensions for long file name support, and discs formatted in the UDF file system.

IsoBuster

IsoBuster is another tool that supports a wide range of discs, and it is available at www.isobuster.com. It recovers data from not only CDs and DVDs but also Blu-ray and HD-DVD discs. It also supports ISO 9660, Joliet, and UDF; will scan for IFO, BUP, and VOB file systems on audio and video DVDs; and supports Mount Rainier CD-RW and DVD+RW discs. IsoBuster is also one of the few tools that provide support for discs created using Macintosh file systems. It has an HFS Reader to support HFS and HFS Plus file systems and a built-in support for Resource Fork extenstions in ISO 9660 and UDF file systems.

CD Data Rescue

CD Data Rescue is a tool developed by Naltech Software (www.naltech.com) that is designed to recover data from damaged, scratched, and defective CD-ROM, CD-R, and CD-RW disks. It supports Mount Rainier/EasyWrite MRW discs, and it can recover data created by CD writing software in ISO and UDF formats.

InDisk Recovery

InDisk Recovery is a tool developed by Octanesoft and is available at www.octanesoft.com. It is designed to recover data from damaged, scratched, defective, or otherwise unreadable CD and DVD discs. It supports the ISO-9660, UDF, and Joliet file systems.

Microsoft Office Repair and Recovery

Data recovery does not just mean restoring deleted files. Occasionally, any previously deleted data may be incomplete or corrupt once it is recovered. Even if a file has not been deleted and restored, the file may be corrupted from improperly shutting down the application or computer, or records within a database may be deleted accidentally. To restore the data within these files, you can use any of the following products:

- OfficeFIX
- Repair My Excel
- Repair My Word

OfficeFIX

OfficeFIX is a suite of products from Cimaware Software (www.cimaware.com) that is designed to repair and recover data from damaged files that were created with Microsoft Office products. The following tools are included in the suite:

- AccessFIX is a tool with features to recover, repair, and undelete Access databases. It recovers not only the data stored in tables but also forms, reports, macros, and other data and elements of the database. It has functions that will restore deleted records from tables and restore password-protected files regardless of whether or not you have the password.
- ExcelFIX is a tool used to recover corrupted Excel files. It will extract the information from a damaged Excel spreadsheet and store it in a new file, complete with any data, formulas, and other content.
- WordFIX is a tool used to recover Word files. This not only includes files created with the PC version of Microsoft Word but also those created with any version of Word for Macintosh. Not only can it recover the text but also higher editions of this tool can recover formatting, tables of contents, embedded images, and other data.
- OutlookFIX is a tool used to repair files used in Microsoft Outlook, and it provides the ability to recover damaged or deleted e-mail, calendars, notes, attachments, and other elements in Outlook. It also provides an inbox repair tool and has the ability to split large files into smaller ones to solve the 2 GB limitation on Personal Folders files (pst).

Repair My Excel and Repair My Word

There are also a number of other products that are designed to repair Microsoft Office products. GetData Software Development (www.getdata.com) provides a number of recovery tools, including the following:

- Repair My Excel (www.repairmyexcel.com) is used to recover spreadsheets created in Excel, including formulas, formatting, and other elements of the file's contents.
- Repair My Word is a free tool available from www.repairmyword.com that is used to recover text from corrupt or damaged Word files. Once recovered, it can then be saved in another Word document.

Compressed Files

Compressed files are files that have had a compression algorithm used to make the files smaller. Because the compressed file may be damaged, any files stored inside of it are inaccessible. This means that before you can recover a Word document, image, or other file, you must first repair the compressed file created with tools such as PKZIP, WinZip, RAR, or other compression software.

Zip Repair

Zip Repair is another tool developed by GetData Software Development (www.getdata.com) and can be downloaded from www.ziprepair.com. It is used to repair corrupted Zip files that have been compressed using WinZip or other software. It supports large file sizes of 2 GB or more, and it can even repair and extract data from spanned Zip volumes that have been split into smaller sets of data. Spanned Zip files are often used to make large Zip files into smaller pieces, allowing them to be e-mailed without worry of any e-mail size limitations.

Deleted Images

Of the various types of data you will try to recover, deleted images are common. If they are personal photographs, losing this data can be particularly devastating since pictures of events and special occasions are something that can never be replaced. Images may be corrupted or deleted from hard disks or directly from the memory card used in a digital camera. At times, the image may have been damaged, and repair of the file is necessary. A number of different tools are available that can be used to recover lost images.

eIMAGE Recovery

eIMAGE Recovery is another tool developed by Octanesoft and is available at www.octanesoft.com. It is designed to recover any digital images or media that may have been lost or deleted from memory cards used by digital cameras. It can restore files from any number of different media, including compact flash, smartmedia, memory sticks, mmd, XD, multimedia, or secure digital memory cards.

Canon RAW File Recovery Software

Canon RAW File Recovery Software (CRW Repair) is another free tool developed by GetData Software Development (www.getdata.com) and can be downloaded at www.crwrepair.com. Canon cameras usually store images in a JPEG format, with RAW images stored inside of a file with a .CRW extension. The CRW file is generally used for processing photos, and it allows a photo's exposure, white balance, and other elements to be manipulated. The CRW file also allows users of the camera to access the JPEG quickly. Unfortunately, because it is a complex file, it can be easily corrupted by such things as a change in file size. CRW Repair will examine the file size to ensure it is correct, and it has the ability to access and extract the JPEG image.

ImageRecall

ImageRecall Software (www.imagerecall.com) provides several editions of software that can be used to retrieve corrupted or deleted files. The *Don't Panic – Photo Edition* allows you to recover pictures and video that have been damaged or deleted from digital cameras.

Using this software, you can obtain and restore data from memory cards, USB storage devices, and other storage media. The tool also provides a thumbnail viewer that allows you to view small images of the pictures available to recover. The software is available only for Windows 2000, XP, and Vista, but support for Macintosh OS X is forthcoming.

RecoverPlus Pro

RecoverPlus Pro is a tool that allows you to recover digital images that may have been damaged or deleted, and it is available at www.arcksoft.com. It provides recovery of a wide variety of graphic formats and will attempt to adjust the image to improve it during recovery. It will attempt to repair any images that are unreadable, including RAW files such as Canon CRW, CR2, and Nikon NEF. It provides a preview pane, which allows you to view a full-size image as well as thumbnails that are rendered from the image file. RecoveryPlus Pro also provides a number of options for scanning and undeleting files, including performing a deep-level scan on damaged media.

Zero Assumption Digital Image Recovery

Zero Assumption Digital Image Recovery was a free stand-alone digital image recovery tool developed by Zero Assumption Recovery (www.z-a-recovery.com). However, the stand-alone version of this tool was discontinued because its features are included in the trial and full versions of ZAR. As part of ZAR, the Digital Image Recovery is designed to restore pictures from digital memory cards. The types of images it can restore are GIF, JPEG, TIFF, CRW, CR2, MOV, and WAV files.

DiskInternals Flash Recovery

As discussed previously, DiskInternals (www.diskinternals.com) provides a number of data recovery solutions including DiskInternals Flash Recovery, which is designed to recover deleted or corrupted pictures from memory cards, including those that have been reformatted or lost due to a hardware malfunction. It provides an easy-to-use interface that allows images to be recovered using a wizard that takes the user step by step through the recovery process.

PC Inspector Smart Recovery

PC Inspector Smart Recovery is a free tool developed by CONVAR that can be used to restore files from memory cards and memory sticks used with digital cameras. It acquires read-only access to the memory card, ensuring that the data is not altered. Once recovered, you can specify where you would like the files to be saved. It is available at www.pcinspector.de/smart_media_recovery/uk/welcome.htm.

Recovering Deleted Partitions

Partitioning a hard disk involves dividing the disk into volumes, which generally appear to the operating system as logical drives, identified by different drive letters. The disk is divided into logical drives for the purposes of performance and organization of the data. Each logical drive can be formatted separately so that each one uses a different file system. Of course, you can partition the disk as a single partition. Partitioning schemes and tools differ depending on the operating system and file system.

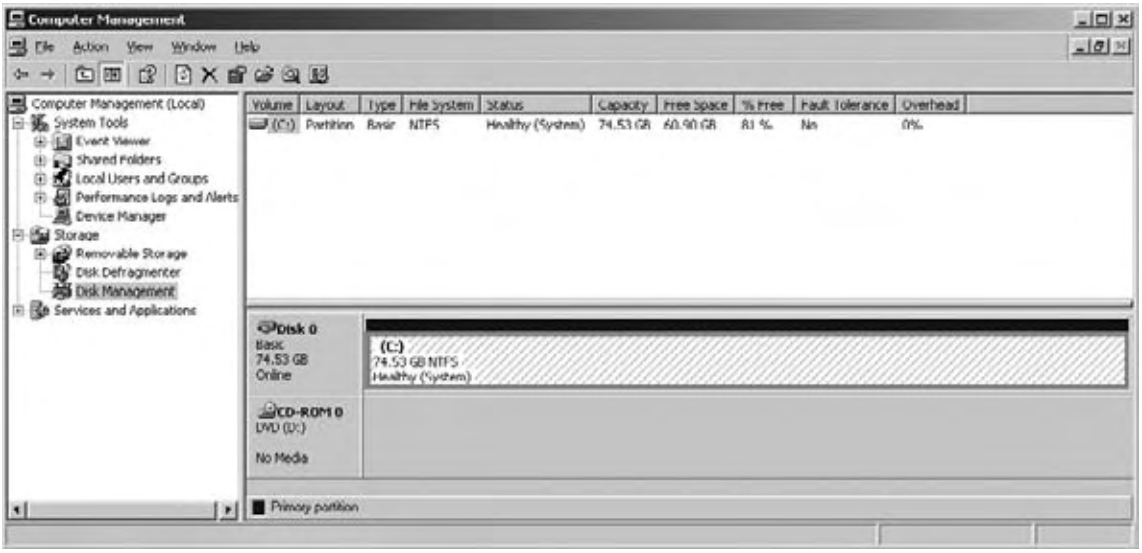
Contrary to popular belief, partitioning utilities do not erase the data on a disk; they only delete and manipulate the partition tables. Although tools such as Partition Magic warn that their use will erase the data on a disk, this is not true; the warning is intended for the average user who will not be able to recover the data after using the utility. Generally, partition utilities will delete the entry for that partition in the partition table so that any space associated with the partition becomes unallocated. Even if the tool overwrites the first sector (sector 0) of the partition before removing it, a backup of the boot sector may still be available and can be restored. Professional data recovery techniques can still recover the data, although the data might be fragmented—that is, the contents of a file could be spread out in different areas of the disk and recoverable in bits and pieces.

Deleting Partitions Using Windows

Partitions can be created and deleted in Windows using the *Computer Management (Local)* console, which is used to control various aspects of your computer. To open the console, you can right-click on the **My Computer** icon on your desktop, and then click the **Manage** menu item on the context menu that appears. Alternatively, you can open the console through the Control Panel. From the Windows **Start** menu, select **Settings** and click the **Control Panel** menu item. When the Control Panel opens, you would then double-click **Administrative Tools**, and then double-click **Computer Management**. It is important to note, however, that to access this tool, you need to use an Administrator account or be a member of the Administrators group.

As shown in Figure 10.8, the Computer Management console provides access to a variety of tools that can be used to manage your computer. By expanding the **Storage** node in the left pane, you can then click on the **Disk Management** item in the console tree. By selecting this item, information about the hard disks and other storage (such as DVD or CD-ROM drives) installed on your computer is shown in the right portion of the screen.

Figure 10.8 Disk Management in the Computer Management Console



By right-clicking on a partition, logical drive, or volume in the lower right pane of the console, a context menu will appear with a **Delete Partition** menu item. If this item is enabled, clicking it will delete the partition. However, the **Delete Partition** menu item will not be enabled in certain circumstances, in which case it is impossible to delete the partition:

- A system volume (which contains files to boot the computer), boot volume (which contains system files), or a volume with an active paging file or crash dump (memory dump)
- An extended partition that is not empty. Before deleting an extended partition, all of the logical drives in that partition must be deleted first.

Deleting Partitions from the Command Line

Partitions can also be deleted from the command line using disk partitioning utilities that require you to type commands from a prompt. There are two commands that may be used on different versions of Windows:

- FDISK
- DISKPART

Using either of these tools, you can view a listing of partitions, determine the number or drive letter of the disk, and delete any existing partitions. It does, however, have the same limitations described previously when we discussed deleting a partition with the Computer Management console. You must be an Administrator or member of the Administrators group, and

you can't delete a system volume, boot volume, or a volume with an active paging file or crash dump. In addition, any logical drives need to be deleted before deleting an extended partition.

FDISK

FDISK is a command line interpreter that is used to create and delete partitions on computers running MS-DOS, Windows 9*x*, Windows NT, or Windows ME. As shown in Table 10.5, there are a number of switches that can be used with this command to view information and perform various actions on a hard disk.

Table 10.5 FDISK Switches

Switch	Description
/MBR	Rewrites the MBR
/CMBR <*disk*>	Re-creates the MBR on a specific disk
/PRI: <*size*>	Creates a primary partition
/EXT: <*size*>	Creates an extended partition
/LOG: <*size*>	Creates a logical drive
/Q	Prevents rebooting the computer automatically after exiting *FDISK*
/STATUS	Shows current status of hard drives
/ACTOK	Forces *FDISK* not to check disk integrity
/FPRMT	Disables prompt for FAT32 support

When typing **FDISK** without any of these switches, the first screen that appears will prompt as to whether you want to enable large disk support. This will determine whether FAT32 is used on the drive. If a hard disk is larger than 512 MB, FAT32 is needed if you want to format a disk over 2 GB as a single drive. Once you press the **N** or **Y** key on your keyboard to indicate whether FAT32 should be used, a menu screen will appear with a number of options:

1. **Create DOS partition or Logical DOS Drive**
2. **Set active partition**
3. **Delete partition or Logical DOS Drive**
4. **Display partition information**
5. **Change current fixed disk drive**, which only appears if you have multiple hard disk drives on your computer

To delete a partition, you would press **3** on your keyboard, which would give you another menu that asks what kind of partition or logical drive you want to delete. The options are

1. **Delete Primary DOS Partition**
2. **Delete Extended DOS Partition**
3. **Delete Logical DOS Drive(s) in the Extended DOS Partition**
4. **Delete Non-DOS Partition**, which is used to delete partitions using NTFS

Upon choosing the type of partition you want to delete, you may be presented with a list of drives that are available to delete. For example, choosing the third menu item will present you with a listing of logical DOS drives in the extended partition. Upon choosing what you will delete, a prompt will appear warning you that all of the data will be lost and asking if you're sure that you want to precede with the deletion. By pressing the **Y** key on your keyboard, the partition or drive is deleted.

DISKPART

DISKPART is a command line interpreter that is used to create and delete partitions on computers running Windows 2000, Windows XP, or Windows Vista. Unlike FDISK, it doesn't provide a series of menus that can be navigated to delete partitions. Instead, as shown in Figure 10.9, typing **DISKPART** at the prompt will activate the command line interpreter, allowing you to enter different commands.

Figure 10.9 DISKPART


Command Prompt - diskpart
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\>diskpart
Microsoft DiskPart version 5.1.3565
Copyright (C) 1999-2003 Microsoft Corporation.
On computer: 06-L2009
DISKPART> list disk
Disk ### Status Size Free Dyn Gpt
Disk 0 Online 75 GB 0 B
DISKPART> select disk 0
Disk 0 is now the selected disk.
DISKPART> list partition
Partition ### Type Size Offset
Partition 1 Primary 75 GB 32 KB
DISKPART>

A number of commands can be used to view information on disks and partitions on your hard disk, as well as perform tasks such as creating and deleting partitions. The commands available through DISKPART are shown in Table 10.6.

Table 10.6 DISKPART Commands

Switch	Description
ADD	Adds a mirror to a simple volume
ACTIVE	Marks the current partition as being the active boot partition
ASSIGN	Assigns a drive letter or mount point to the selected volume
BREAK	Breaks a mirror set
CLEAN	Clears the configuration information or all information off the disk
CONVERT	Converts the disk from one format to another. This will allow you to convert the disk from dynamic to basic, basic to dynamic, MBR to GPT, or GPT to MBR.
CREATE	Creates a volume or partition
DELETE	Deletes a missing disk, selected volume, or a selected partition
DETAIL	Provides details about a disk, partition, or volume
EXIT	Exits the program
EXTEND	Extends a volume
HELP	Prints a listing of Help commands
IMPORT	Imports a disk group
LIST	Prints a list of disks, partitions, or volumes
INACTIVE	Marks the current partition as an inactive partition
ONLINE	Online a disk that is currently marked as offline
REM	Used to comment scripts
REMOVE	Removes a drive letter or mount point
REPAIR	Repair a RAID-5 volume
RESCAN	Forces *DISKPART* to rescan the computer for disks and volumes
RETAIN	Places a retainer partition under a simple volume
SELECT	Moves the focus on an object

Using a combination of these commands, you can delete a partition from a computer, allowing you to later create and format in a particular file format. These steps are outlined in the following exercise, which should not be performed on any production computer.

Deleted Partition Recovery Tools

When a partition is deleted, its entry in the partition table is removed. Although it can appear quite imposing that an entire partition of information is no longer visible, the data hasn't been destroyed from the disk. Essentially, deleting the partition is similar to removing the table of contents from a book; none of the information outside of the table is missing, it just requires other methods to find it. This is where partition recovery tools come into play.

Partition recovery tools perform a number of automated tasks that will attempt to restore a damaged or deleted partition and/or restore data from that partition. The following are some of the automated tasks these tools will use to locate and recover data:

- Determining the error on the disk and allowing the user to choose another partition and make it active
- Scanning the disk space for a partition boot sector or damaged partition information, and then attempting to reconstruct the partition table entry. By finding the partition boot sector, it will have all the information necessary to reconstruct the entry in the partition table. Because both NTFS and FAT32 volumes maintain backup boot sectors, you can recover the volume by restoring the boot sector.
- Scanning the disk space for a partition boot sector or data from deleted partition information, and then attempting to reconstruct the partition table entry

A number of tools are available for partition recovery, each of which has various features that can make it easier to restore data that may have been lost from accidental deletion or damage to the partition. Damaged partitions can occur from power or software failure, the root directory being damaged by a virus, formatting, and from being deleted with tools such as FDISK, DISKPART, or the Disk Management tool in the Computer Management console. Problems can also occur when the partition table, MFT, root directory, or boot record is lost or corrupt. Each of these can cause Windows not to recognize the hard disk, requiring partition recovery tools to be used to recover the partition and data.

Active@ Partition Recovery

Previously in this chapter, we discussed a number of other tools developed by Active@ Data Recovery Software, so it should come as no surprise that this company also has a utility for deleted or damaged partitions. Active@ Partition Recovery is a tool that is available at www.partition-recovery.com, and it will recover FAT12, FAT16, FAT32, and NTFS partitions and logical drives. It also provides the ability to create an image file of the drive to back up data on a disk, as well as back up MBR, partition table, and boot sectors, which can be restored after a problem occurs.

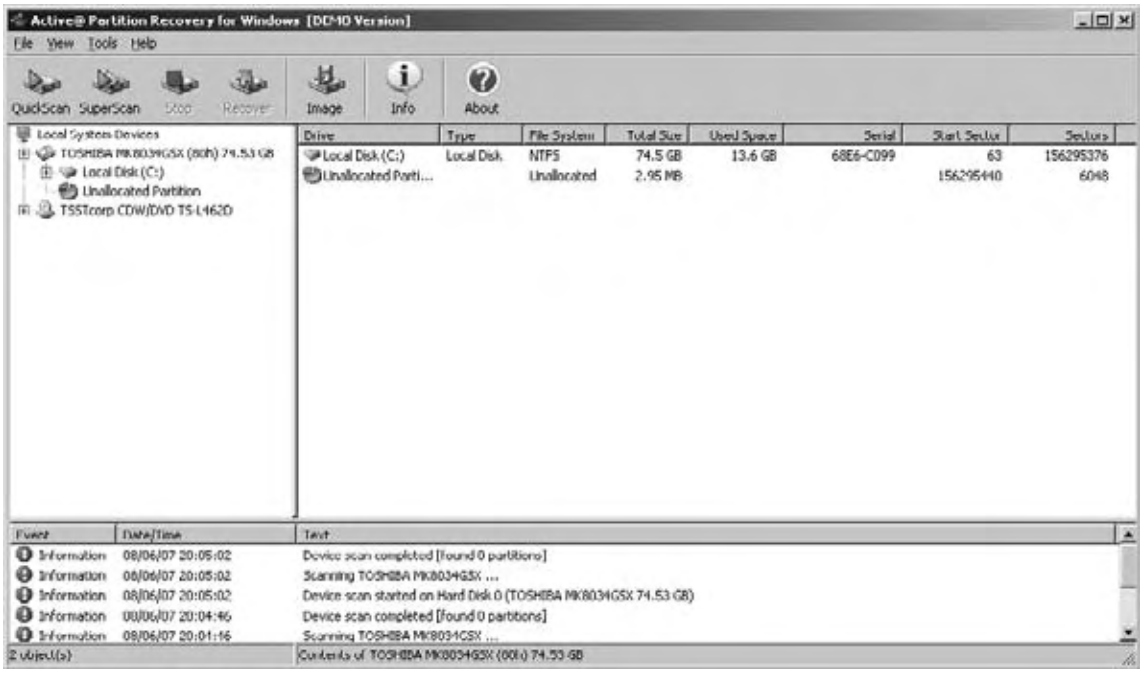
Active@ Partition Recovery will scan the hard disk to locate deleted partitions and logical drive. In doing so, a SuperScan (which does an extensive scan of the hard disk) may even detect partitions after new ones have been created. Even if you've formatted and used these new partitions, the old ones may still be detected. Once detected, you can then undelete primary and extended partitions or logical drives to restore the data. If you're unsure whether the data on the partition is something you want to recover, this tool provides the ability to preview files and folders located on the deleted partition or drive before its recovered.

There are different versions of Active@ Partition Recovery software available to use, including

- Active@ Partition Recovery for DOS, which is a DOS-based version of the tool and is used if the system partition has been lost or if the computer doesn't boot in Windows. It can't recover a partition if it has been overwritten. It is small enough to run from a bootable floppy, and it supports IDE, ATA, and SCSI hard drives.
- Active@ Partition Recovery for Windows, which is the Windows-based version of the tool and is used if a nonsystem partition is lost or when the computer boots in Windows. It can also be used to restore a deleted partition on a USB flash drive or memory card. It supports IDE, ATA, SATA, and SCSI drives.

The Active@ Partition Recovery Professional and Active@ Partition Recovery Enterprise Toolkit both include each of these versions. As shown in Figure 10.10, the interface for the Windows version of this tool is easy to use. By clicking on the **QuickScan** or **SuperScan** buttons, the hard disk is scanned and any partitions are detected. You can preview files and folders detected on the partition and recover the partition. By clicking on the **Image** button, you can create a copy of the logical drive or physical device. Items in the **Tools** menu can also be used to repair the MBR, back up or restore partitioning information, recover the partition, or delete an invalid partition.

Figure 10.10 Active@ Partition Recovery


Active@ Partition Recovery for Windows [DEMO Version]
File View Tools Help
QuickScan SuperScan Stop Recover Image Info About
Local System Devices
TOSHIBA MK8034GSX (80h) 74.53 GB
Local Disk (C:)
Unallocated Partition
TSSTcorp CDW/DVD TS-L462D
Drive Type File System Total Size Used Space Serial Start Sector Sectors
Local Disk (C:) Local Disk NTFS 74.5 GB 13.6 GB 68E6-C099 63 156295376
Unallocated Parti... Unallocated 2.95 MB 156295440 6048
Event Date/Time Text
Information 08/06/07 20:05:02 Device scan completed [found 0 partitions]
Information 08/06/07 20:05:02 Scanning TOSHIBA MK8034GSX ...
Information 08/06/07 20:05:02 Device scan started on Hard Disk 0 (TOSHIBA MK8034GSX 74.53 GB)
Information 08/06/07 20:04:46 Device scan completed [found 0 partitions]
Information 08/06/07 20:01:16 Scanning TOSHIBA MK8034GSX ...
2 object(s)
Contents of TOSHIBA MK8034GSX (001) 74.53 GB

Active@ Disk Image

Whereas Active@ Partition allows you to create a raw image of a logical or physical drive, Active@ Disk Image is a DOS-based tool designed to completely back up and restore an image of your entire hard disk, FAT12, FAT16, FAT32, and NTFS partitions and logical drives. This tool allows you to create both compressed and uncompressed images that contain a mirror of the drive's surface or compressed data images that contain data stored in the clusters. As with other tools from this developer, you can preview the data stored in one or more image files before restoring them. It is available at www.disk-image.net.

DiskInternals Partition Recovery

DiskInternals is another company whose tools were previously discussed. DiskInternals Partition Recovery is a tool that is used to recover deleted and damaged partitions, and it is available at www.diskinternals.com. It supports FAT12, FAT16, FAT32, and NTFS volumes, and it provides features to recover deleted files. Even if files were deleted before the partition was lost, they can still be recovered.

DiskInternals Partition Recovery also provides the ability to create snapshot files of logical drives, allowing any data on a drive to be stored in files that can be kept on another machine until needed. By creating an image of the drive, the data can be restored in case of a problem.

DiskInternals Partition Recovery also provides features to recover the MBR and unformat NTFS and FAT drives. In addition, it can provide recovery for other storage media, such as USB flash drives, memory cards, and other media.

One of the major features of this tool is that it provides a number of wizards that take you step-by-step through the process of recovering partitions. Because it walks you through the process, you can recover a partition without any previous experience.

GetDataBack

GetDataBack is a partition recovery tool developed by RunTime Software, and it is available at www.runtime.org. It will restore data from a variety of sources, including hard disks, memory cards, USB flash drives, iPods, and other media. There are two versions of this software available from this site:

- GetDataBack for NTFS, which is used to recover data from an NTFS file system
- GetDataBack for FAT, which is used to recover data from a FAT file system

Each of these tools can be used on a remote computer, allowing you to connect over a network or serial cable to restore a damaged or deleted partition. This can be useful if you don't want to remove the drive and attach it to another computer or use DOS-based tools to recover the partition.

It provides a wizard-like interface that allows you to specify settings to optimize the recovery process, selecting whether to use default settings, that damage was caused by partitioning software such as FDISK or formatting the drive, or if a new operating system was installed. If there is no problem with the partition, you can also set the tool to simply recover deleted files.

NTFS Deleted Partition Recovery

NTFS Deleted Partition Recovery is a tool that is designed to recover partitions and perform other tasks to restore data that may have been lost. It can retrieve data from formatted, corrupted, or otherwise damaged partitions, as well as from other storage media such as Zip drives, USB flash drives, and memory cards. It is available at www.techddi.com.

Handy Recovery

Handy Recovery is a tool for recovering deleted, damaged, and formatted partitions that is available at www.handyrecovery.com. Using this tool, you can search for files by names or by using a mask, and you can restore entire folders and their contents. You can browse data on the disk, allowing you to see deleted files and folders in addition to ones that have not been deleted. It supports FAT12, FAT16, FAT32, and NTFS file systems and will recover compressed and encrypted files stored on drives formatted as NTFS. It can also recover data

from memory cards and other media. Each file that is displayed provides information on the probability that it can successfully be recovered.

Acronis Recovery Expert

Acronis Recovery Expert is another easy-to-use tool that recovers deleted or lost partitions, and it is available at www.acronis.com. It provides a series of wizards that take you step-by-step through the process of recovering data, and it supports FAT16, FAT32, NTFS, HPFS, Linux ext2, ext3, ReiserFS, and Linux Swap file systems. It also provides features that allow you to work from bootable CDs and floppy disks so you can recover system partitions or systems that fail to boot.

TestDisk

TestDisk is a free tool that can be used to recover data and partitions that have been deleted or lost. It can run on a number of different systems, including Windows NT, Windows 2000, Windows XP, Windows 2003, Linux, FreeBSD, NetBSD, OpenBSD, SunOS, and Macintosh OS X. It has features to fix partition tables, recover FAT32 and NTFS boot sectors from a backup, and rebuild FAT12, FAT16, FAT32, and NTFS boot sectors. It can also locate ext2 and ext3 backup SuperBlock, fix MFT using an MFT mirror, and provide other features that make it possible to recover partitions from a number of different systems. It is available at www.cgsecurity.org.

Scaven

Scaven is a partition recovery tool that can perform unattended multistring searches on hard disks and recover deleted and lost data. It can recover data from accidentally formatted drives, drives with bad sectors, damaged MBRs, and lost partitions. It is available at http://pjwalczak.com/scaven/index.php.

Recover It All!

Recover It All! is a tool developed by DTIData available at www.dtidata.com/recover_it_all.htm. It is designed to restore data lost from accidentally formatting the disk, deleted files, and deleted or damaged partitions and boot sectors. It also provides an executable that will run from a floppy disk allowing you to restore the partition even if the computer's operating system won't start and will prevent overwriting data on the disk by installing the software.

Partition Table Doctor

Partition Table Doctor is a partition recovery tool that checks and repairs the MBR, partition table, and boot sector and recovers damaged or deleted data from FAT16, FAT32, NTFS, ext2, ext3, and Linux Swap partitions on IDE, ATA, SATA, SCSI, and removable hard

disk drives. It provides features to browse the contents of a disk, and it allows you to back up and restore the MBR, partition table, and boot sector. It is available at www.ptdd.com.

Data Acquisition and Duplication

Electronic evidence is fragile by nature and can easily be modified, damaged, or destroyed. Even booting a computer can erase temporary files, modify timestamps, or alter other data in addition to writing data and creating new files to the drive using the boot process. In addition, a computer could be booby trapped so that if a set of keys weren't pressed at bootup or an incorrect password was entered, a program or script could run to reformat the hard disk or overwrite certain data making retrieval of evidence more difficult or impossible. It is because of this that data must be acquired or duplicated from a hard disk before any analysis takes place.

Data acquisition is the act or process of gathering information and evidence. In computer forensics, this means using established methods to acquire data from a suspect computer or storage media to gain insight into a crime or other incident and potentially use it as evidence to convict a suspect. The goal of data acquisition is to preserve evidence, so any tools that are used should not alter the data in any way and should provide an exact duplicate. To prevent contamination, any data that is duplicated should be stored on forensically sterile media, meaning that the disk has no other data on it and has no viruses or defects.

Duplication of data is a critical part of any computer forensic investigation. As discussed in the following sections, a number of tools are available that can create a copy of data. To effectively examine data on a suspect machine, a person performing a forensic examination of the machine needs to create an *image* of the disk. When you create a disk image (a bitstream copy), each physical sector of the disk is copied so that the data is distributed in the same way, and then the image is compressed into a file called an *image file*. This image is exactly like the original, both physically and logically. As an exact duplicate of the data on a suspect machine or storage media, the mirror image includes hidden files, temp files, corrupted files, file fragments, and erased files that have not yet been overwritten. In other words, every binary digit is duplicated exactly.

This is different from other methods that may be used to duplicate data for disaster recovery or other purposes for which every scrap of data isn't necessary. For example, when you use tools to back up a hard disk, the only data duplicated are those that are visible to the file system. Similarly, a disk image may be created as a method of backing up a system, including only files that haven't been deleted. The backup or clone of the disk contents wouldn't include any data that was hidden, deleted, or lost on the disk. If a disaster occurred and/or data needed to be restored, the backup or image could be used to recover data and get systems back online quickly.

From the perspective of someone performing an investigation, backups on magnetic tape or image files stored to disks can be a useful source of evidence. Most organizations will have

a regimen of nightly backups, in which data stored on servers and certain other computers on the network will be copied to magnetic tape or other storage media. Even individuals may back up important data to a disk or location on the network, creating an image of any data that was saved to the disk over a period of time. This can be important to an investigation because it can provide a timeline of what data was saved to certain files at what time. For example, if someone was embezzling from a company, you could see what entries he or she added, deleted, and modified in a spreadsheet or database on a given day by restoring versions of the file from different backups. As discussed later, a number of tools can be used to duplicate data stored on magnetic tapes, and others that can mount an image file created with various types of software. By analyzing a duplicate of the image or backup, you may find important evidence without altering the original media or file.

In an investigation however, most of the time you'll need to use software to create an image of a hard disk, CDs, DVDs, floppy disks, USB flash drives, memory cards from digital cameras, and other storage media. The bitstream copy will contain all of the data stored on the device or removable media and can later be analyzed for potential evidence.

Because you don't boot from the hard disk of a computer you're investigating, you need to connect the computer you're using (with forensic software installed) to the suspect's machine via a network cable or serial cable. Once you've created a physical connection to the computer, you can then run the forensic software and create an image of the machine's hard disk.

Forensic software may include a special program that can run from a bootable floppy or be small enough to fit on a bootable CD or floppy. Upon booting from the CD or floppy, the program on the floppy or CD is run. Depending on the program, it may connect to the forensic software on your computer so that it can then make an exact image of the disk, or it may simply write the image file to your computer.

Alternatively, you could also remove the hard disk from a suspect computer and connect it to your computer or a device that provides write protection. Write protection prevents data from being written to the disk that you're duplicating. When using any forensic software, it is always advisable to use hardware-based write-blocking software. The hardware prevents the forensic software and the operating system you're using from altering data on the suspect hard disk when you attach the hard drive to your computer. Once the hard disk is connected to your computer in this way, you can then begin running the forensic software so that a duplicate of the data is written to an image file on your machine.

Once the image is created, you can then use the forensic software or other tools that we'll discuss in this chapter to mount the image. This allows the software to display the contents of the machine and will allow you to view any data that exists on the machine from which it was acquired. Any software being used should mount the image as a read-only volume so there is no way that any changes can be made to the data. This allows you to view data in areas of the disk, inclusive to partition information, sectors, files stored on

the machine, directory structure, and other information. Because it is read-only, any files you view won't be altered, inclusive to retaining their original timestamps and other attributes.

Because the data you're analyzing is a duplicate of the original source, the original hard disk or storage media is preserved in its original state. If a situation occurred in which the image file or duplicated magnetic media was damaged or destroyed, you could reacquire the data as you did the first time. Of course, any machine used to store forensic image files should be backed up regularly. This will ensure that if a problem does occur that damages or deletes the image file you're working from, you can restore it from an existing backup.

Notes from the Underground…

Booby Traps

Booby trap is military slang for a trap that is set off by a foolish or unsuspecting person. In computers, there are many different kinds of traps that can be set for people whom you do not want to have access. A simple trap is to create an icon on the desktop that someone may want to activate. For example, if a suspect created an icon named "Kiddie Porn," an investigator might be tempted to double-click on it. The trap is that the shortcut actually points to a script or program that will overwrite selected files, encrypt the hard disk, or perform other actions that make analyzing the disk impossible. If you didn't want to use an icon, you could configure your computer to run a particular script or program at startup or shutdown. This could be as simple as adding a batch file to the Windows Startup menu or adding a few lines to an AUTOEXEC.BAT file.

There are many applications on the Internet designed to wipe data from your computer by overwriting deleted files. A simple batch file could delete a folder and then invoke such a program that would overwrite any deleted files with a series of ones and zeros or other data. It could even initiate a low-level format that would erase everything on the disk sector by sector.

The danger of booby traps is another reason why it's very important to analyze a computer using an image of its data. By working from an image file, you won't need to worry about such programs executing and damaging evidence on a machine. The original data remains preserved, and because the forensic software mounts the image as a read-only file, there is no chance that such scripts or programs will modify any data in the image file.

Data Acquisition Tools

Whereas years ago the number of tools that could be used for forensic acquisition and analysis was limited, today there are quite a few options. As mentioned previously, in choosing which tools to use, you should ensure that they do not modify data. You should also evaluate the reputations of these tools and whether they have previously been accepted in court. For example, if you are using a tool such as Encase, which has been widely accepted, there would be less of a chance of it being heavily scrutinized. Since testifying in court can be stressful enough, you don't want to waste time justifying the tools you used to acquire evidence. Although you may still need to validate the tool that was used, using tools that are widely accepted can make the process easier.

In the following sections, we'll discuss a number of data acquisition tools and how their features may be useful. These tools may consist of software used to duplicate data, create image files that may be mounted and analyzed afterward, or hardware-based solutions that can acquire data from a suspect machine. We'll also discuss a number of native commands and tools that can be used to duplicate data on computers and transfer them to another machine.

FTK Imager

The Forensic Tool Kit (FTK) is a product developed by AccessData (www.accessdata.com) and is a fully integrated forensic data acquisition and analysis program. It provides a number of features, such as supporting full text indexing of image files without needing to extract them to a hard disk, and includes a file viewer to preview files. It also has an imaging component (derived from the shareware ISOBuster product) that does a good job of collecting data from CDs and DVDs with any of the commonly supported file systems.

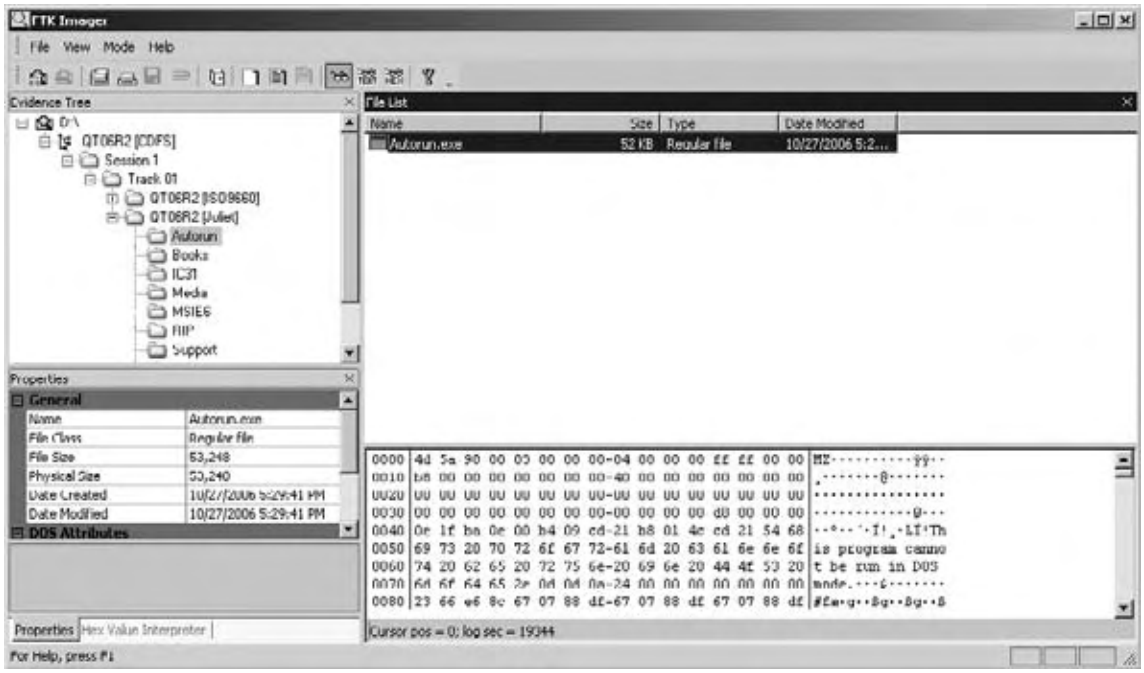
FTK Imager is an imaging tool that allows you to preview data and assess potential evidence on a machine. Using this tool, you can make a forensic image of the data, duplicating everything on the machine so that there is no chance of modifying the original data. By previewing the contents of the image and reviewing the duplicated data, you can then determine if additional analysis is required using FTK.

You can view forensic images of hard disks, floppy disks, CDs, DVDs, and other storage media created with FTK Imager, or you can view images created with other tools. It will read image files created with ICS, SafeBack, and forensic software "forensic software" as meant? and uncompressed images created with Ghost, and it will read or write image files in Encase, dd Raw, SMART, and FTK image formats. This means that even if another organization or person with different software created a forensic image, you could still view the image file and determine if there was any evidence on media. This is particularly useful in a situation such as when an internal investigation was conducted, a forensic image was created from a suspect computer, and police now need to view the evidence that was acquired.

In addition to the image file formats that can be made for analyzing disks, there are also a number of file formats that can be read and created for CD and DVD forensics, including ISOBuster CUE, CloneCD, Alcohol, PlexTools, and Virtual CD.

As shown in Figure 10.11, FTK Imager provides an easy-to-use interface. Once an evidence file is opened, you can view the folder structure in the **Evidence Tree**, located in the upper left pane. By selecting a folder, you can then view files stored in that folder in the **File List,** located in the upper right pane. To preview a particular file, you can select it in the upper pane and view an image of pictures, hexadecimal data, text, and previews of other data in the lower right pane. Additional information on the file, including any DOS attributes the file might have, can be viewed in the **Properties** pane in the lower left side of the screen.

Figure 10.11 FTK Imager



Now that we've discussed a number of the features of this product, let's try it out and create a forensic image using the tool. Instructions for how to create an image file are outlined in Exercise 4.3.

SafeBack

SafeBack was one of the earliest DOS-based tools for acquiring evidence from a computer sector by sector. It was developed by NTI and is available at www.forensics-intl.com/safeback.html. SafeBack has been marketed to law enforcement agencies since 1990 and has been used by the FBI and the Criminal Investigation Division of the IRS to create image files for forensics examination and evidentiary purposes. Booting from a floppy disk, it can make a duplicate image of everything on a hard disk, preserving its integrity and allowing you to analyze evidence without worry of modifying it.

SafeBack is capable of duplicating individual partitions or entire disks of virtually any size, and the image files can be transferred to SCSI tape units or almost any other magnetic storage media. The product contains CRC functions to check integrity of the copies and date and timestamps to maintain an audit trail of the software's operations. The vendor provides a 3-day computer forensics course to train forensics specialists in the use of the software. (In fact, the company does not provide technical support to individuals who have not undergone this training.) Although it is a DOS-based tool, it can be used to copy DOS, Windows, and UNIX disks (including Windows NT/2000 RAID drives) on Intel-compatible systems. Images can be saved as multiple files for storage on CDs or other small-capacity media. To avoid legal concerns about possible alteration, no compression or translation is used in creating the image.

DriveSpy

DriveSpy is another DOS-based data acquisition tool developed by Digital Intelligence Forensic Solutions and is available at www.digitalintelligence.com. If you are familiar with DOS commands, it is an easy-to-use tool because it uses a DOS prompt and many of the commands used in DOS to navigate through the directory structure of a system being analyzed. In addition to these commands, special commands are used to invoke features of DriveSpy.

Although it is a DOS-based tool, it can acquire evidence from partitions using FAT, non-DOS partitions, and hidden DOS partitions. This not only includes files that are visible to the file system but also data that has been deleted or that exists in slack space and unallocated space on the disk. It supports acquisition from hard drives larger than 8.4 GB as well as floppy disks and other storage media. It provides a built-in sector and cluster Hex Viewer to view data, and it has the ability to create and restore compressed forensic images of partitions.

A beneficial feature of this tool to an investigation is its logging capabilities. You can configure DriveSpy to document how the acquisition of evidence is conducted, including logging every keystroke that you made. These are written to a log file and can be used later for inclusion in a report of what actions you took to acquire the evidence.

Mount Image Pro

Mount Image Pro is a tool developed by GetData Software Development (www.getdata.com) that allows you to mount and view forensic image files created by EnCase, Unix and Linux DD images, SMART, and ISO images of CDs and DVDs. It is used by a variety of organizations, including the FBI, state and federal law enforcement, and various other investigative agencies and firms. Once an image is mapped to a drive letter, you can explore the data acquired through the image and use third-party tools discussed previously (such as Recover My Files) to restore files.

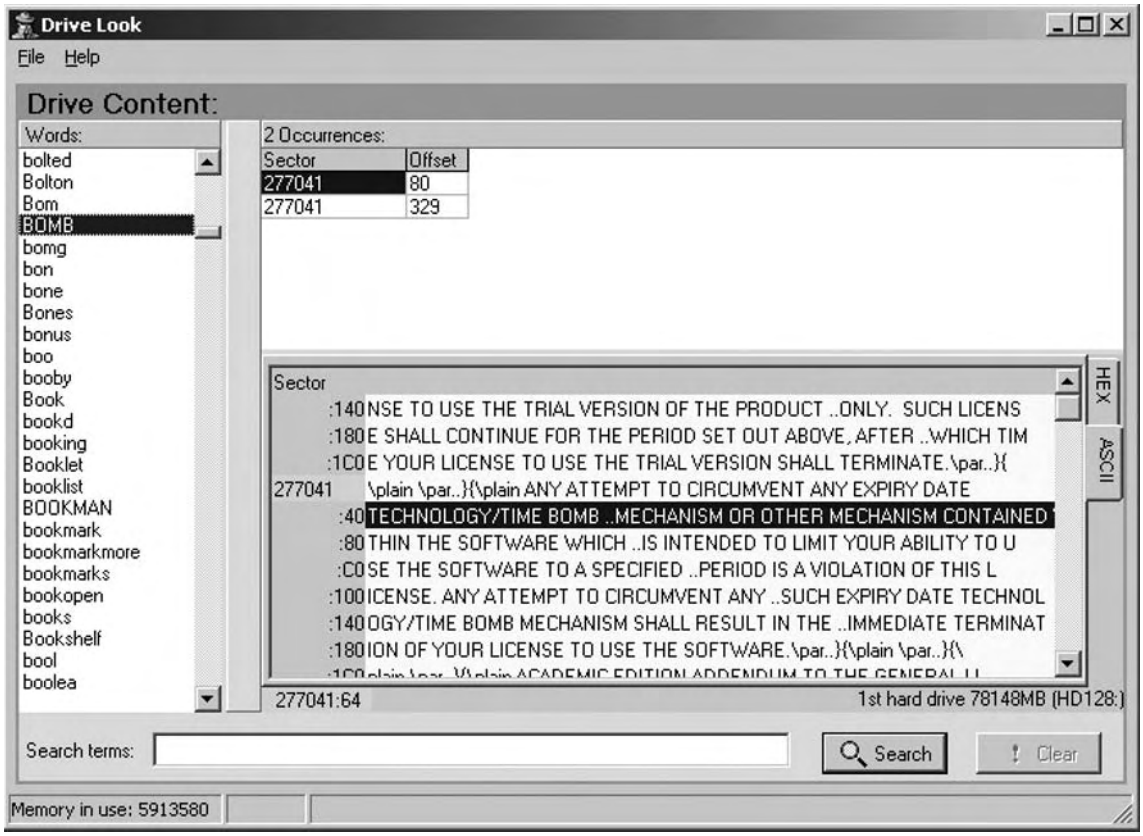
Because Mount Image Pro can mount different types of forensic image files, you don't need additional copies of software or dongles used by software such as EnCase to view evidence acquired in the image. This can save a considerable amount of money because machines used to analyze the image don't require the original software used to acquire the evidence. Additionally, if an EnCase image file is password protected, you can still open it even if you don't have the password. It is available at www.mountimage.com.

DriveLook

DriveLook is a free forensic search tool developed by Runtime Software and is available at www.runtime.org/drivelook.htm. It is designed to index all of the text that was written to a hard disk or other media. The drives searched can be physical or logical drives, or remote drives to which you connect using a serial cable or network connection using TCP/IP. You can also use this tool with image files, such as those created with Runtime's DiskExplorer or GetDataBack, allowing you to search for words that were stored on a suspect machine. Once it has indexed the words on the media and saved them to a table, you can search by keywords or browse the table to view the location of where these words are stored.

When DriveLook first opens, it provides an easy-to-use wizard that takes you through the steps of indexing a drive. When it first starts, you configure where you'd like to store a log file, words to include in the search and other settings, or accept the defaults before DriveLook scans the drive. Once the scan is finished, you click the **Next** button to begin indexing the drive. Once complete, clicking the **Next** button to view the screen shown in Figure 10.12. As shown in this figure, by selecting a word from the left pane, you can view the sector in which it is located on the drive in the upper right pane and view how it appears in the lower right pane. The interface also allows you to filter the results using a search field on the bottom of the screen.

Figure 10.12 DriveLook



Although it is free, when you initially install DriveLook, it installs as an evaluation copy. Before using it, you need to click in the **Help** menu, and then click on the **Register** menu item. Upon doing so, a dialog box appears in which you enter the following information:

- Name: Runtime Software
- License key: BKZSBKFTUPYEK

Upon clicking **OK**, the program is then registered, and you can use it as a fully licensed version.

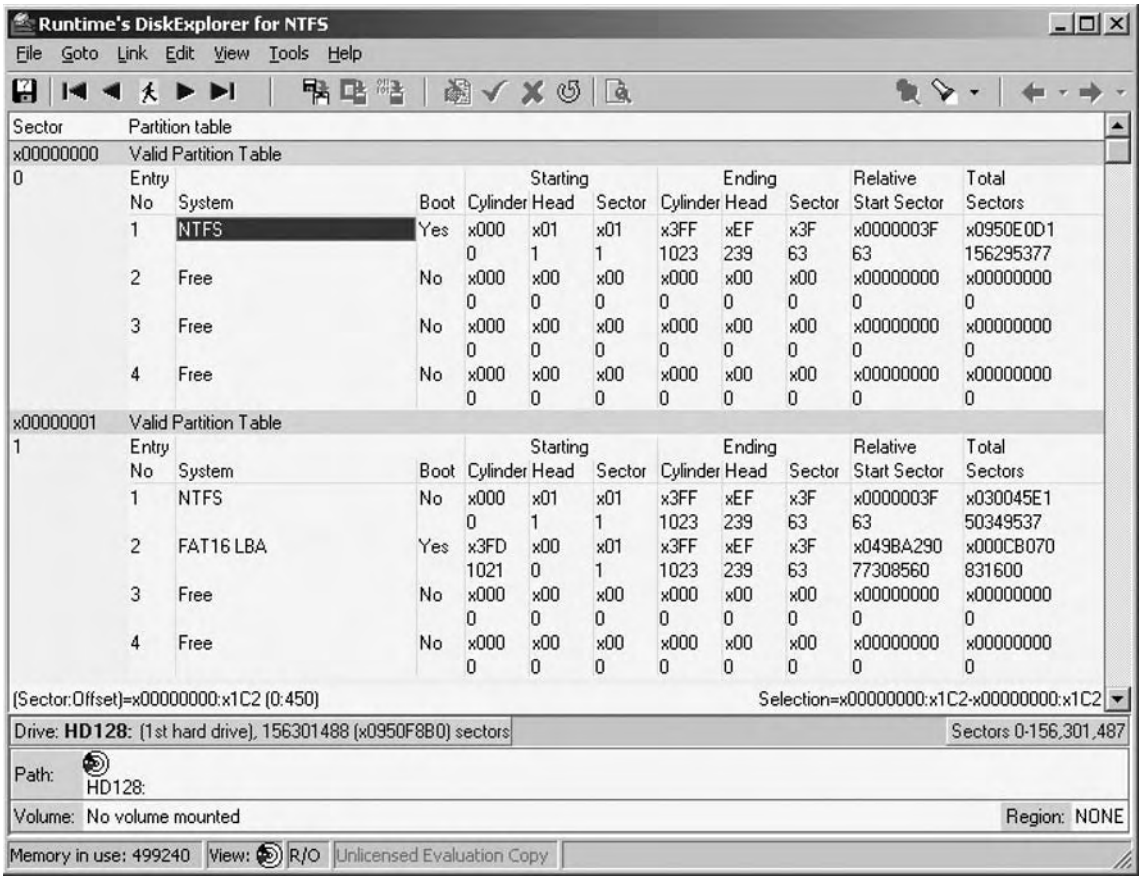
DiskExplorer

DiskExplorer is another tool that allows you to browse the contents of a hard disk. Developed by Runtime Software and available at www.runtime.org/diskexpl.htm, there are two versions of this tool available to use:

- DiskExplorer for NTFS
- DiskExplorer for FAT

Both of them are disk editors, which are used to respectively browse NTFS and FAT file systems and recover data stored on a disk. They allow you to view the contents of a physical or logical drive or image files. As shown in Figure 10.13, when viewing information on a disk, you can analyze the partition table, MFT, boot record, and index buffers. Using back and forward buttons, you can navigate through data on the disk, or you can use items on the menu to go to a particular sector or area of the disk to view its data. DiskExplorer also features search capabilities that allow you to search for text, view files and their properties, and identify to which file a cluster belongs. Using this tool, you can edit sectors of the disk and also recover files that may have been lost or deleted.

Figure 10.13 DiskExplorer


Runtime's DiskExplorer for NTFS
File Goto Link Edit View Tools Help
Sector Partition table
x00000000 Valid Partition Table
0 Entry No System Boot Starting Cylinder Starting Head Starting Sector Ending Cylinder Ending Head Ending Sector Relative Start Sector Total Sectors
1 NTFS Yes x000 x01 x01 x3FF xEF x3F x0000003F x0950E0D1
0 1 1 1023 239 63 63 156295377
2 Free No x000 x00 x00 x000 x00 x00 x00000000 x00000000
0 0 0 0 0 0 0 0
3 Free No x000 x00 x00 x000 x00 x00 x00000000 x00000000
0 0 0 0 0 0 0 0
4 Free No x000 x00 x00 x000 x00 x00 x00000000 x00000000
0 0 0 0 0 0 0 0
x00000001 Valid Partition Table
1 Entry No System Boot Starting Cylinder Starting Head Starting Sector Ending Cylinder Ending Head Ending Sector Relative Start Sector Total Sectors
1 NTFS No x000 x01 x01 x3FF xEF x3F x0000003F x030045E1
0 1 1 1023 239 63 63 50349537
2 FAT16 LBA Yes x3FD x00 x01 x3FF xEF x3F x049BA290 x000CB070
1021 0 1 1023 239 63 77308560 831600
3 Free No x000 x00 x00 x000 x00 x00 x00000000 x00000000
0 0 0 0 0 0 0 0
4 Free No x000 x00 x00 x000 x00 x00 x00000000 x00000000
0 0 0 0 0 0 0 0
(Sector:Offset)=x00000000:x1C2 (0:450)
Selection=x00000000:x1C2-x00000000:x1C2
Drive: HD128: (1st hard drive), 156301488 (x0950F8B0) sectors
Sectors 0-156,301,487
Path: HD128:
Volume: No volume mounted
Region: NONE
Memory in use: 499240
View: R/O
Unlicensed Evaluation Copy

This tool can also duplicate data on a hard drive by creating an image file. After creating an image file, you can preserve data on the disk for later analysis. Once the image file is created, you can then open it again with DiskExplorer and analyze its contents or use other tools to restore it on another machine.

SnapBack DatArrest

DatArrest is a suite of data duplication and forensic tools developed by SnapBack and is available at www.datarrest.com. One of the benefits of this tool is that data can be acquired from a machine using a program on a bootable floppy. Images created not only include files that can be seen by the file system but also those that have been deleted, encrypted, or that are in slack space on the disk. Even CMOS settings used by the computer are captured as part of the image. Using the modules and utilities included with this product, you can acquire data and create images to and from hard disks, removable media, and magnetic tapes.

SCSIPAK

SCSIPAK is a suite of tools developed by Vogon and is available at www.vogon.co.uk/data-conversion/scsipak.htm. It provides data recovery and data conversion between Windows NT4 or Windows 2000 and other systems such as UNIX, DEC, ICL, and IBM Mainframe. When attempting to view a backup tape created by one of these systems on a Windows NT or 2000 tape drive, the Windows machine may be unable to read it and view it as blank. SCSIPAK extends the capabilities of the Windows tape drive so that this data is now readable. You can download data from tapes or optical discs and write them to up to seven drives simultaneously. Using SCSIPAK, you can copy tapes, transfer data from one system to Windows NT or Windows 2000, or transfer data between disk, tape, or optical disc.

IBM DFSMSdss

Data Set Services (DFSMSdss) is a utility developed by IBM that is designed for disaster recovery and management of data. Developed as part of the Data Facility Storage Management Subsystem (DFSMS), DFSMSdss is used to move or copy data between different types of storage media so that storage on various servers can be managed and space won't fill up on one particular volume of a server. It also allows you to back up and restore data and copy backups to other storage media, regardless of whether it is between the same or different types of media. It is available at www-03.ibm.com/systems/storage/software/sms/dss/index.html.

Hardware Tools

Although to this point we have discussed how software may be used for data acquisition, this is not the only method of duplicating data. Using these tools in a forensic lab or in the field, you can create forensic images that can later be analyzed for potential evidence. In some cases, these tools serve as portable forensic labs, allowing you to perform acquisition and analysis before the computer is even removed from the crime scene.

A number of these tools will store data directly on a hard disk inside of the device, provide the ability to transfer image files from the device to another computer in a forensic lab, and can then wipe the device's drive to make it forensically sterile. This ensures that there is no contamination from other data on the drive each time data is acquired from a new machine.

ImageMASSter Solo-3 Forensic

ImageMASSter Solo-3 Forensic is a hardware tool developed by Intelligent Computer Solutions and is available at www.icsforensic.com. It is a portable, hand-held device that can acquire data from suspect machines at speeds exceeding 4 GB per minute. By connecting the hard disks directly to the machine using a drive-to-drive interface or external Firewire/USB interface, you can create an exact replica of data from one or two drives simultaneously without any reduction in speed. Specifically designed for computer forensic purposes, it can acquire data from IDE, SATA, and SCSI hard drives.

LinkMASSter-2 Forensic

LinkMASSter-2 Forensic is another hardware tool developed by Intelligent Computer Solutions and is also available at www.icsforensic.com. The device attaches to a computer through a Firewire or USB port and can then create an image of any data on the machine. Software is used to boot the machine and connect to the LinkMASSter, allowing it to acquire data using the Firewire or USB ports. Data on the original machine is protected during acquisition due to write-blocking features in the device.

ImageMASSter 6007SAS

ImageMASSter 6007SAS is a powerful tool for creating images of data from suspect machines, and it is a useful part of any forensic lab. It was developed by Intelligent Computer Solutions and is available at www.icsforensic.com. It duplicates IDE, SAS, SATA, and IDE hard drives, and it can migrate server data from SCSI to SAS/SATA. ImageMASSter can also acquire data from multiple hard disks and store multiple images on one hard drive. It is the only duplication system on the market that supports SAS (Serial Attach SCSI) hard drives and can copy multiple drives simultaneously at high speeds. It also includes a 1 GB network connection that can be used to transfer files to and from a network drive. To acquire and analyze data, the system provides a Windows XP-based interface that allows you to copy data from Windows, Macintosh, and UNIX file systems.

RoadMASSter-3

Whereas many tools are developed for use inside of a forensic lab, the RoadMASSter-3 is a data acquisition and analysis tool designed for use in the field. The RoadMASSter-3 is developed by Intelligent Computer Solutions and is available at www.icsforensic.com. It is designed to create an image and analyze data acquired from suspect hard drives. It can connect to an unopened computer using Firewire and USB ports, and it can connect directly to IDE, SATA, SAS, and SCSI hard drives. It can also acquire data from multiple drives to a single target drive, making acquisition faster.

The RoadMASSter-3 also allows you to analyze data immediately. It has a 15-inch color display in its case that allows you to view the data stored in the image file so that you can determine whether any evidence exists on the machine while you're still in the field.

Disk Jockey IT

Disk Jockey IT is a portable, hand-held hardware tool developed by Diskology and is available at www.diskology.com. It is the smallest write-blocking and disk copy device available for computer forensics. It can be used as a write-blocking device for data acquired using Firewire and USB connections to a suspect computer. The device can then be connected to a Macintosh or Windows computer in write protect mode to analyze the data without fear of altering the data.

Backing Up and Duplicating Data

Although the previously discussed tools are used for investigations in which there is a need to acquire an exact duplicate of data for the purpose of analyzing and potentially using certain data as evidence in court, this is not the only reason why data is duplicated. Many tools are used to create backups of data so that systems can be restored in the event of a virus or other malicious software, catastrophic failure of a hard disk, intrusion, or other events that could result in the corruption, loss, or deletion of data.

Images are also made of computers when there is a mass deployment of a system. For example, on a network, most of the workstations will have the same operating systems, applications, and other software installed and configured the same way. By setting up one computer and making an image of it, you can then restore the image on other computers being deployed. You would only need to change the computer's name, IP address, and a few other settings on the machine to make it unique on the network. Because the operating system and other software are only installed once, and every other computer has an image restored on it, this saves time and money in setting up machines being deployed on the network.

Backups and duplicates of data are particularly useful during investigations. Because a computer may be seized during an investigation, a new computer can quickly be deployed with a clean image restored to it. This minimizes the impact of losing a computer for any length of time because a replacement computer can be deployed in its place almost immediately.

R-Drive Image

R-Drive Image is a data duplication tool developed by r-Tools Technology and is available at www.drive-image.com. It is designed for backing up and duplicating data, and it creates a byte-by-byte copy of data on hard drives, partitions, and logical disks. These files can then be stored on another hard disk, network drive, or other storage media and restored as needed.

Save-N-Sync

Save-N-Sync is a tool designed by Peer Software to synchronize data stored in a directory on a laptop or desktop computer with data stored on another storage device or location, such as a network server. The data can be backed up to a location you specify, or you can synchronize it so that any changes on either the source or the target folder will be reflected in both directories. You can manually or automatically have the data on a computer backed up or synchronized, ensuring that if a problem occurs, the data can be restored from another location. It is available at www.savensync.com.

The standard version of this tool allows you to choose a single directory to synchronize with another directory, whereas the corporate version allows up to 15 directories to be synchronized. Although this may sound like a small number of directories, you can select for all subdirectories below each directory to be included in the synchronization process. For example, if you choose your C: as the source folder, and had any subdirectories included in the synchronization, then your entire C: drive would be synchronized.

QuickCopy

QuickCopy is a tape duplication system developed by Shaffstall Corporation that is designed to make tape-to-tape copies of backups. When data from a server or other computer is backed up to magnetic tape, copies may be required for offsite storage. When a nightly backup is made, QuickCopy can be used to make a duplicate of this data, providing a backup of the backup tape.

Although useful in disaster recovery situations, it can also be useful in making duplicates of tapes so they can be analyzed for potential evidence. QuickCopy verifies the data byte-for-byte and also provides the ability to copy a single tape to an image file that can be stored on a hard disk. It is available at www.shaffstall.com/products/software/qctsw/.

Acquiring Data in Linux

In addition to the data recovery and forensic tools that run or can be used on machines running Linux, UNIX, and other Posix operating systems, certain commands can be used on these operating systems to copy data on a machine and transfer it across a network. Versions of these same tools are also available for Windows operating systems but have their roots in Linux/UNIX machines. These commands are

- *dd*, which is a convert and copy command
- *Netcat*, which is used to transfer data and for other functions

DD

The *dd* command is a common UNIX command that is also used in Linux; it is used to convert and copy data. A version of DD is also available for systems running Windows, which is available at either http://unxutils.sourceforge.net or www.cygwin.com. It allows

you to copy a hard disk to another disk drive or magnetic tape, or vice versa. In doing so, data is transferred byte-for-byte, creating an exact mirror image. The syntax for using this command is

```
dd <options>
```

As you can see from this syntax, there are options used with this command to specify different aspects of converting and copying data. For example, if you wanted to copy the contents of one disk to another, you would use the command with the *if* (input file) and *of* (output file) options:

```
dd if=/dev/hda of=/dev/hdy
```

Similarly, if you were going to back up the disk to an image on a hard disk, you could use the following:

```
dd if=/dev/hda of=/path/to/image
```

Additional options can be used with the *dd* command to specify the size of blocks to copy, whether data is to be converted when copying, and other options shown in Table 10.7.

Table 10.7 *dd* Options

Option	Description
if=inputfile	Specifies where to input data from (file or device)
of=outputfile	Specifies where to output data (file or device)
ibs=bytes	Number of bytes to read at a time
obs=bytes	Number of bytes to write at a time
bs=bytes	Number of bytes to read and write. This is used instead of ibs and obs, and it specifies the same number of bytes to use for both input and output.
bs=bytes	Number of bytes to convert at a time
skip=blocks	Specifies to skip blocks in the input file before copying
seek=blocks	Specifies to skip blocks in the output file before copying
count=blocks	Copy blocks from the input file, instead of everything at the end of the file
conv=conversion	Specifies to convert the input file before copying to an output file. Conversion methods include ■ ascii, which converts EBCDIC to ASCII ■ ebcdic, which converts ASCII to EBCDIC

Continued

Table 10.7 Continued

Option	Description
	■ ibm, which converts ASCII to alternate EBCDIC ■ block, which replaces the input newline with a padding of spaces to fit the size of cbs ■ unblock, which replaces trailing space characters in datasets of size cbs with newline characters ■ lcase, which converts uppercase characters to lowercase ■ ucase, which converts lowercase characters to uppercase ■ swab, which swaps every pair of input bytes ■ noerror, which ignores read errors ■ notrunc, which specifies not to truncate the output file ■ sync, which pads every input block to the size of ibs with null bytes if its shorter than the specified size

Using the *dd* command, you can even extract the MBR from a disk. Remember that the MBR and partition table are located in the first bytes of the disk. As such, by specifying to only acquire the first few bytes of data from the disk, you can acquire a backup of this data. For example, if we were going to acquire the file from the disk we previously used as an example, we would use the following command:

```
dd if=/dev/hda of=/path/to/image count=1 bs=512
```

To later restore this data to the disk in the event of a problem, we could reverse the input and output file information and instruct dd to put the data back:

```
dd if=/path/to/image of=/dev/had
```

Netcat

Netcat is a tool that is available on all Posix operating systems, including Linux and UNIX, used to transfer data across a network using Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). A version of Netcat is also available to use on systems running Windows 9*x*, Windows NT, Windows ME, Windows 2000, and Windows XP. Using this tool, you can create an image of the partition and send it to a machine or transfer files between remote computers. It also provides port scanning capabilities and the ability to connect via Telnet.

Netcat is run in server mode on one machine and client mode on another machine to communicate with one another. Netcat in server mode listens on a port you specify and sends data to the client that connects to that port number. To run Netcat from the server, the following syntax would be used, where *port number* is the port that you want the server to listen on and that clients will connect to:

```
nc -l <-options> <port number>
```

From the client, you would use the following syntax, where *ip address* is the IP address of the server, *options* are the options you can use, and *port number* is the port that the client will use to connect to the server:

```
nc <-options> <server ip address> <port number>
```

As shown in Table 10.8, Netcat provides a number of additional options that can be used for listening on a network and connecting to computers.

Table 10.8 Netcat Options

Option	Description
-d	Background mode. Detach from console
-e *program*	Inbound program to execute
-g	Source routing flags
-h	Help
-i *seconds*	Delay interval for lines sent, ports scanned
-l	Listen for inbound connections
-n	Numeric only IP addresses (no DNS)
-o	Hex dump of traffic to a file
-p	Local TCP/IP port to listen to. This is used in server mode with –l.
-r	Random local and remote ports
-s *addr*	Local source address
-t	Answer TELNET negotiation
-u	Use UDP to listen on port. This is used in server mode with –l.
-v	Verbose. –v-v will put it into ultra verbose mode.
-w *seconds*	Timeout for connects and final net reads, before Netcat will automatically quit
-z	No I/O mode. Used for scanning ports

Netcat can also be used with other tools to create a compressed image of the disk and allow others to download the file. In the following example, the *dd* command is used to create an image of a disk, with gzip used to compress it. Netcat is then used to listen to port number, making the file available:

```
dd if=/dev/hdb | gzip -9 | nc -l 4321
```

Damage & Defense…

Netcat as a Hacking Tool

Netcat is a useful tool for simple file transfers and sending backups of data across a network, but it is also referred to as the "hacker's Swiss army knife." It is a powerful tool, and in the wrong hands it could be devastating to a system. In fact, Symantec lists Netcat as a hacking tool on its Web site and provides information on how to remove it from a system (www.symantec.com/security_response/writeup.jsp?docid=2005-120115-5633-99). It is an incredibly useful and possibly dangerous tool on networks. Although network administrators commonly use it for debugging and transferring files, it could be used by a hacker to obtain unauthorized access to a Web site, execute programs remotely, and transfer data from a machine to an intruder's computer.

Summary

As discussed in this chapter, there are a number of ways in which data may be deleted from a system. Operating systems provide commands to remove data from a hard disk or other media, such as menu items or commands used from a command prompt. When a file is deleted using the **DEL** key or commands in compliant programs, it will generally be sent to the Recycle Bin. When this occurs, the user can recover the file him- or herself. If the file has been emptied from the Recycle Bin or deleted using other methods, then you may need to use tools to recover the file. Even if an entire partition is deleted, and the volume is formatted, this doesn't mean that the data is gone. As we discussed, data recovery and computer forensic tools may still be able to retrieve the data from a system.

In this chapter, we saw that computer forensics requires specialized tools to duplicate data from a hard disk or other media. In some cases, the tools may be software, whereas hardware may also be used to acquire evidence from suspect machines. The hardware and software-based solutions will create bitstream images of the disk and include any data that is visible to the file system, as well as hidden, deleted, and fragments of files, the MBR, partition table, and any number of other items on a disk. Software can then mount these images and view the data in a variety of ways. Many tools will even allow you to preview the data without extracting it from the forensic image file.

Frequently Asked Questions

Q: I have restored a deleted file, but when I open it, the file is empty. What happened?

A: The file is probably zero length. Because the file size is zero, an empty file is created. If clusters were used by other files, it was overwritten. Because so much data was overwritten by another file, there was nothing to recover.

Q: I am preparing to back up data on my hard disk using a partition recovery tool. What data should I back up?

A: You should make an image of all data on the disk so that this can be restored in the event of a problem. The image file could be stored on a network server or another disk, where it will be safe until you need it. You should also back up the MBR, partition table, and boot sectors, which can be restored after a problem occurs.

Q: If software is available to create a forensic image of a hard disk or other media, what is the benefit of forensic hardware?

A: Different types of forensic hardware are available. Write-blocking tools may be used to prevent data from being written to a suspect hard disk or other media. Forensic tools are also available, which can be connected to an unopened computer or hard disk and acquire data from it. In some cases, they even provide the ability to perform analysis of the acquired data in the field. Many of these tools are portable, allowing you to carry them to a crime scene, and can provide the ability to acquire data from multiple hard drives at high speeds.

Q: Why is it important that all the software used by law enforcement officers be licensed and registered? Law enforcement budgets are often tight; why not use freeware as much as possible?

A: Some freeware and shareware tools available on the Internet are good tools, and the price is certainly right. However, there are some dangers in using these programs for forensic purposes. First, you never know exactly what you're getting when you download a free program (and you certainly can't ask for your money back if it doesn't work properly). Downloads can be infected with viruses or Trojans that can damage the systems on which you use them. Using unlicensed software (illegal copies) is even worse. The opposing attorney(s) will have a field day if he or she discovers that the police used pirated or "borrowed" software in the investigation. This behavior can destroy the credibility of the people who conducted the forensic examination and even result in losing the case. In addition, with properly purchased and registered software, you will be able to get technical support from the vendor if necessary. Makers of computer forensics software often offer discounts to law enforcement agencies, making it easier to afford the proper tools for the job. After all, officers and agencies probably wouldn't suggest saving money by buying their duty weapons from a pawnshop; that's because these are essential tools of the trade and must be as reliable as possible. For the cybercrime investigator or technician, the same is true of the forensics software that is used to collect and preserve evidence that can make or break a criminal case.

Chapter 11

Forensic Discovery and Analysis Using BackTrack

Solutions in this chapter:

- **Digital Forensics**
- **Acquiring Images**
- **Forensic Analysis**
- **File Carving**
- **Case Studies: Digital Forensics with the BackTrack Distribution**

☑ **Summary**

Introduction

Computer forensics is an emerging discipline that has exploded in popularity in recent years. It is an exciting field and a lot of security professionals are interested in learning more about it. Although computer forensics traditionally is not the first thing that comes to mind when people think about doing penetration testing, a large number of companies are starting to turn to computer forensic tools because of the amount of technology they use.

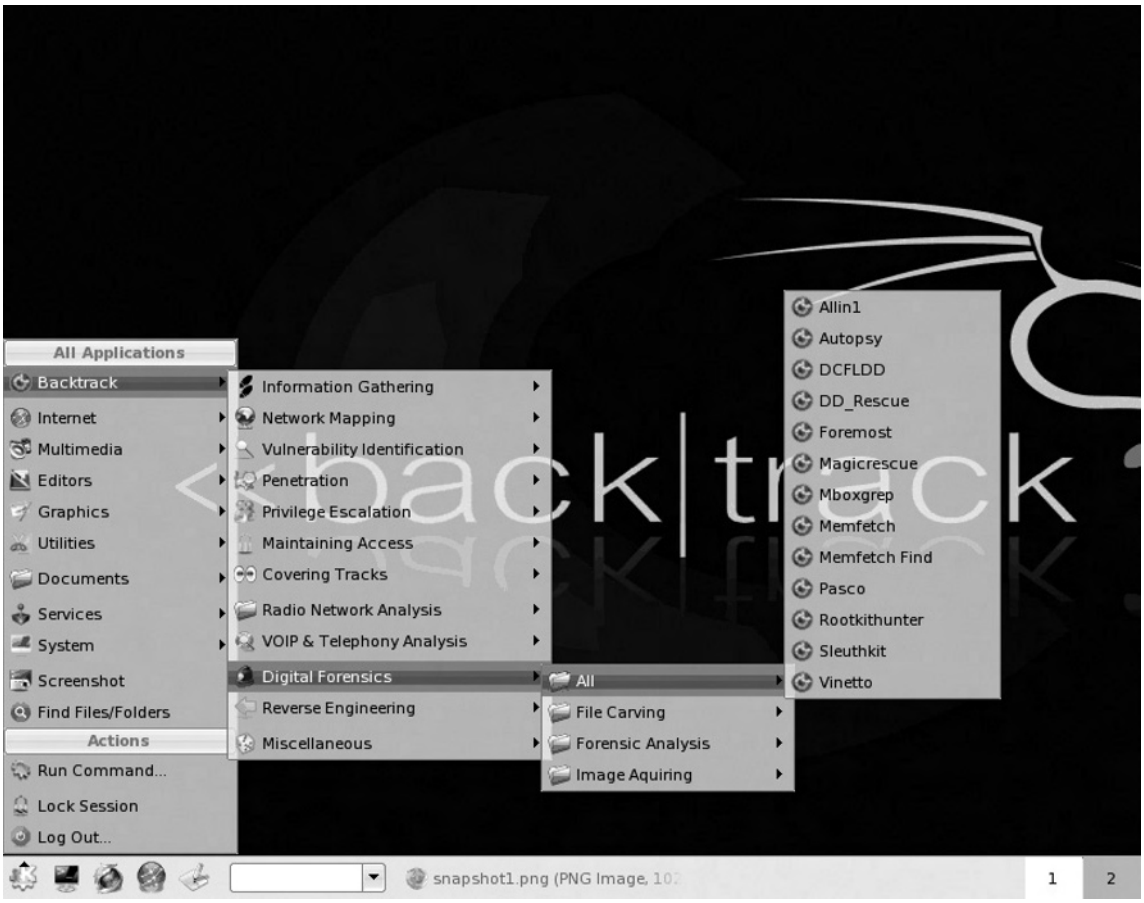
As the field of computer forensics continues to expand, the number of forensic hardware and software tools has also continued to increase. Some of the forensic software packages released by the leading vendors tend to very expensive; single licenses for some forensic software can cost as much as $4,000. For this reason, open source forensic tools can be extremely useful for companies or individuals with limited financial means. One such tool, Auditor (the predecessor to BackTrack), was limited to four forensic tools. The four forensic tools included with the Auditor Linux distribution, as shown in Figure 11.1, were Autopsy, Recovered, Testdisk, and Wipe. The Auditor Live CD was mainly used for penetration testing, so it had limited forensic capabilities.

Figure 11.1 The Four Forensic Tools in the Auditor Security Collection



BackTrack was also designed primarily as a penetration testing tool, but it includes 13 very powerful open source forensic tools, as shown in Figure 11.2.

Figure 11.2 The 13 Forensic Tools in the BackTrack Security Collection



The BackTrack tools are divided into three categories: Image Acquiring, File Carving, and Forensic Analysis. The tools in the Image Acquiring category allow you to acquire bit-by-bit images of drives. The two File Carving tools, Foremost and Magicrescue, assist with the file recovery process. The nine tools in the Forensic Analysis category allow you to search media for items such as e-mail messages, browser history, and thumbnails.

As you can see, computer forensic tools are becoming a necessary part of computer security distributions such as BackTrack. The fact that the number of forensic tools has increased significantly in the BackTrack distribution also may indicate that computer

forensics is now considered an important component of penetration testing. We'll be discussing how to use BackTrack throughout this chapter. You can download the BackTrack ISO from http://www.remote-exploit.org/.

Digital Forensics

Most people today are using some type of electronic device, be it a cell phone, global positioning system (GPS), personal digital assistant (PDA), laptop PC, desktop PC, TiVo system, or game console. Most of these devices leave behind artifacts or footprints that can provide a wealth of information about how they were used. Similarly, many companies provide employees with Blackberrys, laptops, cell phones, and access to workstations and servers, and employees often leave behind digital artifacts regarding how they used the devices. Digital forensics is the discipline of retrieving artifacts such as e-mail messages or Internet browser cache from media so that those items can be analyzed and processed. As the number of electronic devices continues to increase in our everyday lives, the need for additional digital forensic tools will continue to increase.

Digital or computer forensics is a relatively new field in information technology. The U.S. Department of Defense (DoD) started the Defense Cyber Investigations Training Academy in 1998 to deal with the ever-increasing number of computer-based crimes. Many colleges and universities have started offering both bachelor's and master's degrees in computer forensics. Chaplain College in Burlington, Vermont, and Villa Julie College in Stevenson, Maryland, are among the schools offering such programs. In addition, a number of companies are releasing proprietary tools that will help examiners conduct investigations.

A variety of open source forensic tools, some of which are included in BackTrack, have been developed to perform some of the same functions as their closed source counterparts. Computer forensic software tends to very expensive; sometimes single licenses can exceed $4,000. For universities, colleges, or businesses that need multiple licenses, the cost of computer forensic software represents a major barrier. Distributions such as Helix and BackTrack, along with open source tools, allow individuals with limited budgets to perform acquisitions and forensic analysis.

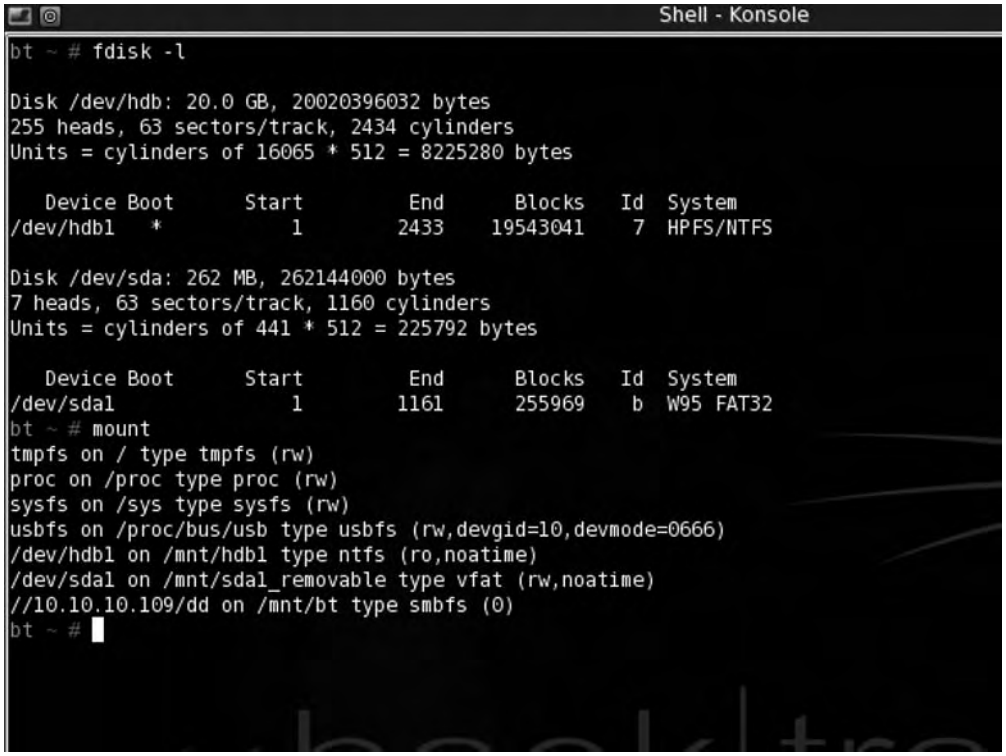
Acquiring Images

Even though tools are included in BackTrack for acquiring images, the use of a BackTrack Live CD with default settings is not considered forensically sound. Most of the commonly used forensic distributions auto-mount drives as read-only; by default, BackTrack mounts drives as read/write, as shown in Figure 11.3. However, BackTrack mounts New Technology File System (NTFS) drives as read-only. Computer forensic examiners do not mount drives as read/write because a savvy attorney could argue that the investigator altered the drives to incriminate his client. Sloppy forensic work could indeed help a defense attorney place doubt in the minds of a jury. Other Live CDs used primarily as forensic tools, such as Helix, mount drives as read-only.

NOTE

If you are going to use BackTrack to acquire images, you need to use the *bt nohd* parameter at boot time so that BackTrack does not auto-mount drives as read/write. Even after you use the *bt nohd* parameter at boot time, you must mount the suspect's drive as read-only. It is critical that computer forensic investigators use the *bt nohd* switch every time at boot. If you are going to use BackTrack frequently for forensic purposes, you should modify the distribution so that the *bt nohd* option is the default. The *bt nohd* helps forensic investigators ensure that the data they are copying does not become contaminated. From a practical standpoint, if you are not conducting actual forensic investigations and are not very comfortable with mounting drives in Linux, using the *bt nohd* will add an extra layer of difficulty.

Figure 11.3 Not Forensically Sound! BackTrack Mounting Drive As Read/Write by Default


Shell - Konsole
bt ~ # fdisk -l
Disk /dev/hdb: 20.0 GB, 20020396032 bytes
255 heads, 63 sectors/track, 2434 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hdb1 * 1 2433 19543041 7 HPFS/NTFS
Disk /dev/sda: 262 MB, 262144000 bytes
7 heads, 63 sectors/track, 1160 cylinders
Units = cylinders of 441 * 512 = 225792 bytes
Device Boot Start End Blocks Id System
/dev/sda1 1 1161 255969 b W95 FAT32
bt ~ # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/hdb1 on /mnt/hdb1 type ntfs (ro,noatime)
/dev/sda1 on /mnt/sda1_removable type vfat (rw,noatime)
//10.10.10.109/dd on /mnt/bt type smbfs (0)
bt ~ #

Damage & Defense …

The Source Disk: Only Read-Only, Never Read/Write

You should not use BackTrack to do forensic acquisitions unless you use the *bt nohd* parameter at boot so that BackTrack does not auto-mount drives as read/write. If you used the normal BackTrack Live CD to do an acquisition and the suspect's drive was mounted as read/write, the evidence would be contaminated.

The BackTrack images acquiring menu lists two tools, dd_rescue and dcfldd. Both of these tools are variants of the dd program whose origins you can trace to the early days of UNIX. The dd program itself is also included in the distribution, so there are actually three tools within BackTrack that you can utilize for image acquisition. The dd program is a powerful forensic tool which will allow you to make a bit-by-bit copy of a drive. Many of the forensic tools found in the industry are just front ends for the dd program. The dcfldd program is newer and has some additional features that are useful for forensic investigators. The dd_rescue program will assist users with recovering files from a disk that has been damaged.

Linux dd

Both the dd_rescue and dcfldd programs are based on the original UNIX dd program. The Linux dd program will allow users to make a bit-by-bit copy of media. The dd program has been around for a long time; it was first seen in the UNIX operating system in the 1970s. The normal copy command, *cp*, just copies files or folders. The *dd* command actually copies the 512-byte data blocks. A computer forensic investigator can use the dd program to create an exact copy of a suspect's media. The investigator can use the md5sum program in BackTrack to compare the suspect's original drive to the copied drive. If the md5sum hashes are the same, a jury can be confident that the data sets are the same; the chance of two files having the same MD5 hash is 1 in 2^{128}. You can open and examine dd images using Autopsy or proprietary tools such as EnCase, Forensic Tool Kit (FTK), or X-Ways Forensics.

An investigator will work with only the bitstream copy of the suspect's media. Conducting an investigation on the original media would cause that media to be contaminated. The bit-by-bit copy will include everything on the original media, including deleted files, folders, and slack space. When a user deletes files on a system, some file

systems do not delete the actual data on the disk. Instead, the space on the disk is marked as available in the file table. Unless those clusters are overwritten, that data can still be present on the disk. Bitstream copies of media made using dd will contain deleted files and folders that have not been overwritten. In summary, when you use dd to make a bitstream copy, all the data from a suspect's media, including deleted files, folders, and slack space, will be acquired.

Notes from the Underground…

Slack Space

A wealth of forensic information can often be retrieved from slack space. Slack space is the unused portion of a cluster on a disk. If a cluster is 4k or 4,096 bytes, and a file is 4,097 bytes, that file will take up two clusters on the hard drive and the remaining 4,095 bytes will be called slack space. It is possible that slack space could contain valuable data from previous files occupied in a cluster.

If you want to use the BackTrack CD to make a bitstream copy of a hard drive using the dd utility, it would be wise to carry a USB 2.0 device with you that is large enough to allow you to copy an entire disk from a computer. You can also copy your image to a FireWire device, but not all computers that you are acquiring disks from will have FireWire ports. If you don't have a portable storage device, you can send the image to another machine over the network using netcat. Bit-by-bit copies can take awhile, so you want to ensure that you have ample time to copy drives.

WARNING

Before we discuss the steps to acquire a drive, it is imperative that you understand the drive letter assignments and partitioning scheme to effectively copy someone else's media. A wrong command on certain systems could result in deleting data or contaminating evidence. Please refer to Tables 11.1, 11.2, 11.3, and 11.4 for a guide on how Linux systems handle drive lettering and partition numbering.

Table 11.1 Linux IDE Hard-Disk Naming Conventions

HDA	Primary master IDE
HDB	Primary slave IDE
HDC	Secondary master IDE
HDD	Secondary slave IDE

Table 11.2 Linux SCSI, USB, and SATA Hard-Disk Naming Conventions

SDA	First SCSI (or USB or SATA) device
SDB	Second SCSI (or USB or SATA) device
SDE	Fifth (or USB or SATA) device
SDZ	Twenty-sixth (or USB or SATA) device

Table 11.3 Linux Partition and Logical Drive Numbering

1–4	Partitions, primary or extended
5 and up	Logical drives

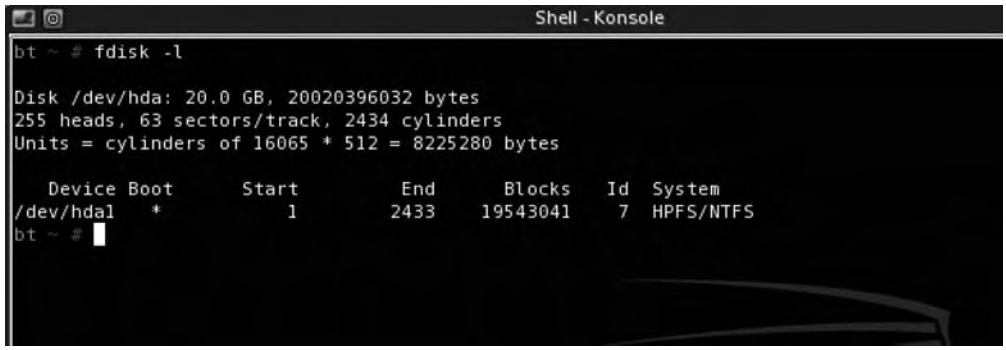
Table 11.4 Examples of Linux Naming Conventions

HDA3	Primary IDE master, third partition
HDD4	Secondary IDE slave, fourth partition
HDC5	Secondary IDE master, first logical partition
SDB9	Second SCSI device, fifth logical partition

If you are just trying to copy a drive and you are not concerned about forensically sound procedures, boot the target system to your BackTrack CD or USB thumb drive. If you do not want to contaminate the media, use the *bt nohd* switch at boot. You will need to become comfortable with the way Linux gives numbers and letters to disks and partitions. Typing **fdisk -l** in the terminal window will give you information about the disks and partitions in the machine.

Figure 11.4 displays the output of the *fdisk* command with a single IDE drive in the target system.

Figure 11.4 The fdisk –l Command Displaying the Partitions


Shell - Konsole
bt ~ # fdisk -l
Disk /dev/hda: 20.0 GB, 20020396032 bytes
255 heads, 63 sectors/track, 2434 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 2433 19543041 7 HPFS/NTFS
bt ~ #

Most of the systems you come across will probably have a single hard drive, as in Figure 11.4. If the computer is running Windows 2000, XP, 2003, or Vista, there is a good possibility that the system has an NTFS partition. Default versions of BackTrack will mount any NTFS partitions as read-only. This is good news from a technical standpoint, because if a drive is mounted as read-only, contamination will not occur. However, from a forensic standpoint, because BackTrack mounts some drives as read/write by default, you need to use the *bt nohd* switch at boot, use a write blocker, or use a different distribution altogether.

Connect your USB (or FireWire) mass storage device to the computer that contains the drive you want to acquire. If the default BackTrack settings are used, you will receive a pop-up box about new media being detected; just select **Open** in a new window, and click **OK**. If the *bt nohd* cheat code is used at boot, you will need to mount this device as read/write. The destination or target disk is the one disk that forensic examiners mount as read/write. Note that no matter what switch is used at boot, when a drive is mounted from the command line, it will be mounted as read/write unless you use the *–o ro* switch with the *mount* command. Your USB mass storage device should have a file allocation table (FAT32) partition. (Most drives come from the factory with FAT32 file systems.) Type **fdisk -l** at the terminal and you will notice that the USB mass storage device is listed; most likely as /dev/sda1, as in Figure 11.5.

Figure 11.5 The fdisk –l Command Displaying the Added USB Drive

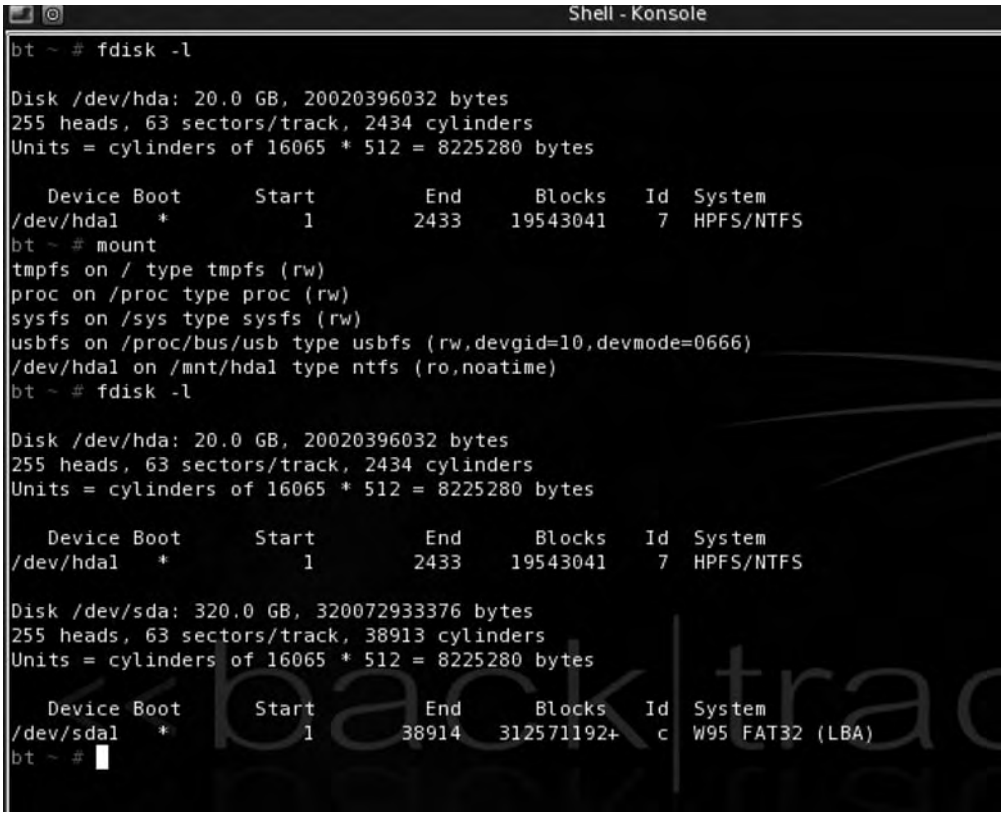

Shell - Konsole
bt ~ # fdisk -l
Disk /dev/hda: 20.0 GB, 20020396032 bytes
255 heads, 63 sectors/track, 2434 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 2433 19543041 7 HPFS/NTFS
bt ~ # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/hda1 on /mnt/hda1 type ntfs (ro,noatime)
bt ~ # fdisk -l
Disk /dev/hda: 20.0 GB, 20020396032 bytes
255 heads, 63 sectors/track, 2434 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 2433 19543041 7 HPFS/NTFS
Disk /dev/sda: 320.0 GB, 320072933376 bytes
255 heads, 63 sectors/track, 38913 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 38914 3125711192+ c W95 FAT32 (LBA)
bt ~ #

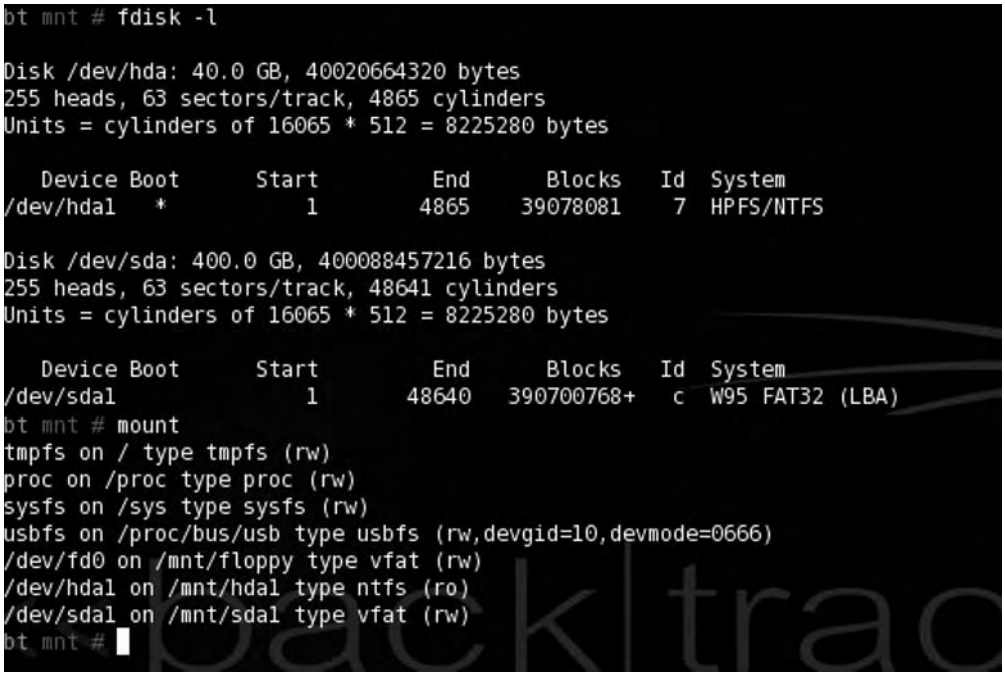
Figure 11.6 shows how the *mount* command is used to verify that the USB device is mounted as read/write. You can use the *dd* command to create a bit-by-bit image of the system's internal hard drive. The dd image created will need to be split because the maximum file size allowed on a FAT32 partition is one byte less than 4 GB. A good rule of thumb is to keep file sizes slightly smaller than what would fit on a CD-ROM. Some other proprietary imaging software splits evidence files into separate 640 MB files by default. You can divide the size of the evidence file into 640 MB chunks by using the *split* command. The *split* command adds a two-letter alphabetical extension to the image filename. The naming convention goes from aa to az, ba to bz, up to za to zz. The split images are treated as a single file when they are imported into computer forensic analysis software.

Notes from the Underground…

dd Images

I recently came across an article about a "sting" in which some people planted pictures and movies on a PC and took the PC to a store (I will not say the name) to get software installed. That store had more than one location, so they took the PC to several locations. They wanted to see whether store employees would access certain files and folders on the PC, so they enabled auditing. To their surprise, many of the stores viewed and copied these files, and posted the logs on the Internet. If the employees made a forensic copy of the drive and completed only the task the people had asked them to complete on the actual machine, the people would have had no way of knowing the employees had "copied" or viewed their data. Warning: Do not make a forensic copy of a customer's data without the customer's knowledge, because doing so is illegal and unethical.

Figure 11.6 The mount Command Showing the Added FAT32 USB Drive with Read/Write Permissions


bt mnt # fdisk -l
Disk /dev/hda: 40.0 GB, 40020664320 bytes
255 heads, 63 sectors/track, 4865 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 4865 39078081 7 HPFS/NTFS
Disk /dev/sda: 400.0 GB, 400088457216 bytes
255 heads, 63 sectors/track, 48641 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sda1 1 48640 390700768+ c W95 FAT32 (LBA)
bt mnt # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/fd0 on /mnt/floppy type vfat (rw)
/dev/hda1 on /mnt/hda1 type ntfs (ro)
/dev/sda1 on /mnt/sda1 type vfat (rw)
bt mnt #

Configuring & Implementing…

Imaging a Hard Drive with dd Using BackTrack

1. Turn the computer on and immediately enter the BIOS.
2. Verify that the CD/DVD-ROM drive is the first device listed in the boot order.
3. Insert the BackTrack CD and boot to BackTrack with the *bt nohd* option at the **boot:** prompt.
4. Log in as root with a password of toor, and type **startx** to bring up the GUI.
5. Open a terminal session.
6. Type **fdisk –l**. You will see the drive letter and numbering designation of the source disk. (This is most likely HDA1 for a single IDE drive with one partition.)
7. Mount the drive as read only:

```
ex:
cd /mnt
mkdir hda1
mount –o ro /dev/hda1 /mnt/hda1
```

   Optional: If you want to verify the integrity of the drive you are about to copy, you will need to do an *md5* on the drive (as seen in Figure 11.7); for example, **md5sum /dev/had**. (Notice here that we use the drive, not the partition.)

8. Plug in the USB mass storage device.
9. Type **fdisk –l** again. You should see a new, added drive letter and partition, most likely sda1. This is your destination disk. Mount this drive as read/write:

```
ex:
cd /mnt
mkdir sda1
mount -o rw /dev/sda1 /mnt/sda1
```

10. Just to verify that the device is mounted properly as read/write, type **mount**.

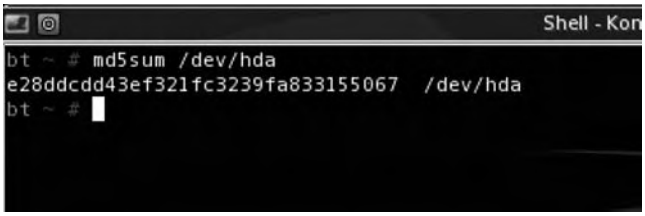
Continued

11. **cd /mnt/sda1** (or whatever the designation for your destination disk was after step 9).
12. **mkdir BackTrack**
13. **cd BackTrack**
14. **dd if=/dev/hda bs=512 conv=sync,noerror | split -b 640m – hd.img**. This process could take awhile depending on the speed of your devices (see Figure 11.8). After this is finished, you have your exact copy which you can examine with the Autopsy tool. If you want to verify the forensic integrity of the dd copy, you will need to perform step 14.

    Optional: **cat hd*.img | md5sum**. This process, seen in Figure 11.8, could take awhile. If you see the same hash value that you saw in step 7, you can be confident that the drives have the same data sets.

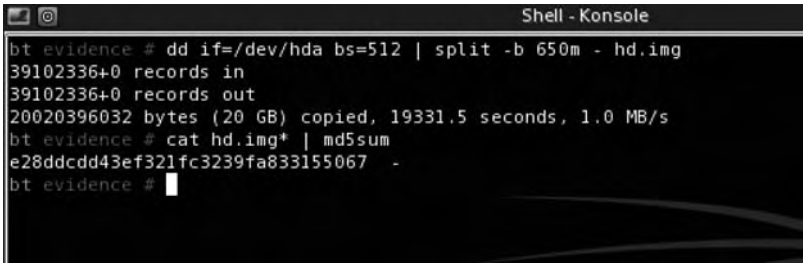
If you are interested in verifying that your image has the same bits as the original, you can use a hash. Hashing the original drive is a relatively simple process that you can accomplish by using the *md5sum* command. In Figure 11.7, we used *md5sum* to create a hash value from the entire drive.

Figure 11.7 Using the md5sum Command to Get the Hash Value of the Hard Disk



You can hash all of the dd files by using the *cat* command and piping the results into the *md5sum* command, as seen in Figure 11.8.

Figure 11.8 Using the md5sum Command to Get the Hash Value of the Combined Image Files


Shell - Konsole
bt evidence # dd if=/dev/hda bs=512 | split -b 650m - hd.img
39102336+0 records in
39102336+0 records out
20020396032 bytes (20 GB) copied, 19331.5 seconds, 1.0 MB/s
bt evidence # cat hd.img* | md5sum
e28ddcdd43ef321fc3239fa833155067 -
bt evidence #

NOTE

The MD5 hash from Figures 11.7 and 11.8 is a match. The fact that these two data sets have the same MD5 value means that the copy is forensically equivalent to the original. The MD5 hash is an extremely accurate measure; the chance of two files having the same MD5 hash is 1 in 2^{128}. You can also use other hash values, such as SHA1, on both data sets. The sha1 and md5 tools are also part of the Sleuth Kit suite of tools included with BackTrack.

Sometimes a mass storage drive is unavailable, but there is plenty of space on your network server. If this is the case, you can boot the system to BackTrack and redirect the dd output to the server using netcat. The media on the local and target systems will have the same MD5 hash value if the procedure is carried out properly.

Notes from the Underground…

dd Images and netcat

If a hacker can get network access to a system, he can use the *dd* command in conjunction with netcat to copy the system's entire drive and redirect the output to his own system. If he performs this process on a live system, taking an MD5 hash is unnecessary because the hash value will constantly change on a live system. Warning: Do not make a forensic copy of someone's data without her knowledge because doing so is illegal and unethical.

Configuring & Implementing…

Imaging a Hard Drive with dd Using netcat

1. Turn the computer on and immediately enter the BIOS.
2. Verify that the CD/DVD-ROM drive is the first device listed in the boot order.

Continued

3. Insert the BackTrack CD and boot to BackTrack with the *bt nohd* switch.
4. Log in as root with a password of toor, and type **startx** to bring up the GUI.
5. Open a terminal session.
6. Type **fdisk –l**. You will see the drive letter and numbering designation of the source disk. (This is most likely HDA1 for a single IDE drive with one partition.)
7. Mount the drive as read-only:

```
ex:
cd /mnt
mkdir hda1
mount –o ro /dev/hda1 /mnt/hda1
```

   Optional: If you want to verify the integrity of the drive you are about to copy, you will need to do an *md5* on the drive (see Figure 11.7):

```
md5sum /dev/had
```

   (Notice here that we use the drive, not the partition.)

8. On the server to which you want to transfer the dd image, navigate via the command line to the directory on the destination partition. Start the netcat listener by typing **nc –v –w 30 –l –p 2007 > hd.img**.

   Note: Most servers have drivers with either NTFS or EXT2/3 partitions, so you will not have to be concerned about a 4 GB file size limit and you will not need to use the *split* command. The image created on the server will be a single file. Make sure the firewall has that port open, and on the off chance that you use Windows servers click **Unblock** if you see the pop-up box shown in Figure 11.9.

9. On the machine to be imaged, type **dd if=/dev/hda bs=512 conv=sync,noerror | nc –w 2 <IPAdressofServer> 2007**. (Note: I picked a random port of 2007. The ports should match on both machines.) This process could take awhile, depending on the speed of your devices. After this is finished, you will have your exact copy, which you can examine with the Autopsy tool. If you want to verify the forensic integrity of the dd copy, you will need to perform step 10.
10. **md5sum hd.img** If you are using a Windows system as your server, you will need the md5sum tool. You can download this file for Windows from www.etree.org/md5com.html. (If you do download that file, put it in your system32 directory.) This process could take awhile. If you see the same hash value that you saw in step 7, you can be confident that the drives have the same data sets.

Figure 11.9 Windows Firewall Alert Stating That You Are Opening a netcat Listener

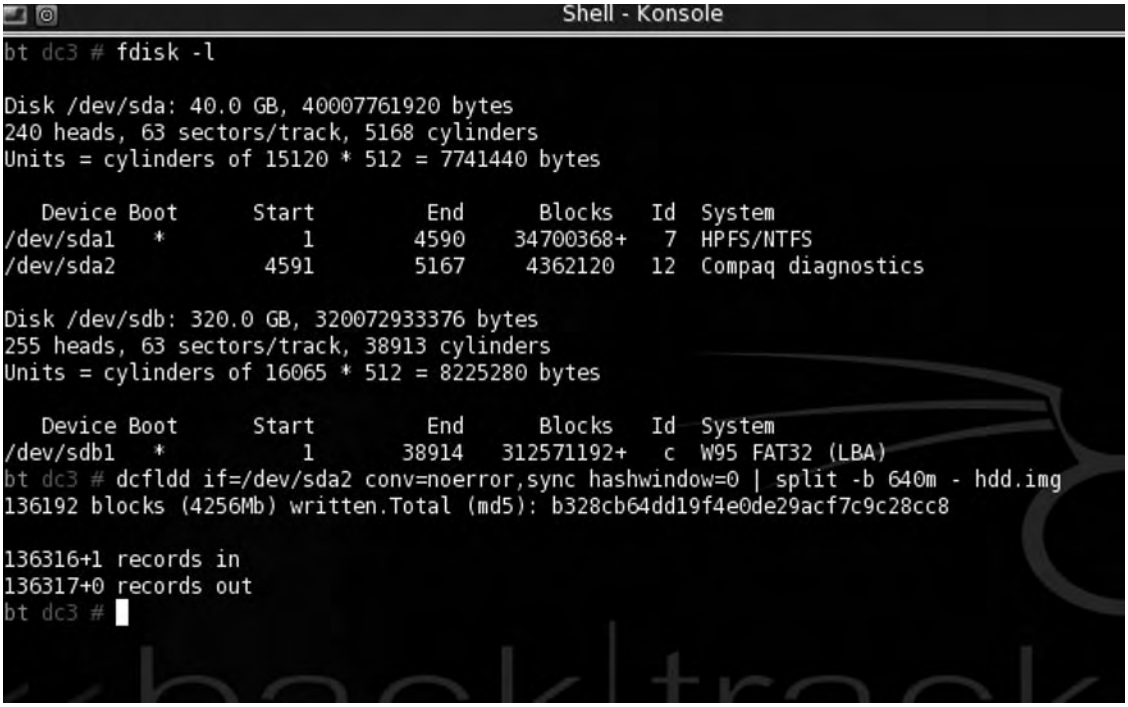

Windows Security Alert
To help protect your computer, Windows Firewall has blocked some features of this program.
Do you want to keep blocking this program?
Name: nc
Publisher: Unknown
Keep Blocking
Unblock
Ask Me Later
Windows Firewall has blocked this program from accepting connections from the Internet or a network. If you recognize the program or trust the publisher, you can unblock it. When should I unblock a program?

Linux dcfldd

The second tool you can use in BackTrack to acquire images is dcfldd. This tool was developed by Nick Harbour when he worked for General Dynamics at the Defense Cyber Forensics Lab (DCFL) in Maryland. DCFL is one of three main entities of the DC3 DoD complex in Maryland. The DC3 Cyber Crime Center was created in October 2001 to help DoD agencies solve an increasing number of computer crimes. The other two parts of DC3 are the Defense Cyber Investigations Training Academy and the Defense Cyber Crime Institute. The Defense Cyber Investigations Training Academy offers a large number of computer forensic and network security classes. The Defense Cyber Crime Institute is primarily responsible for creating accepted standards and techniques for conducting computer forensic investigations.

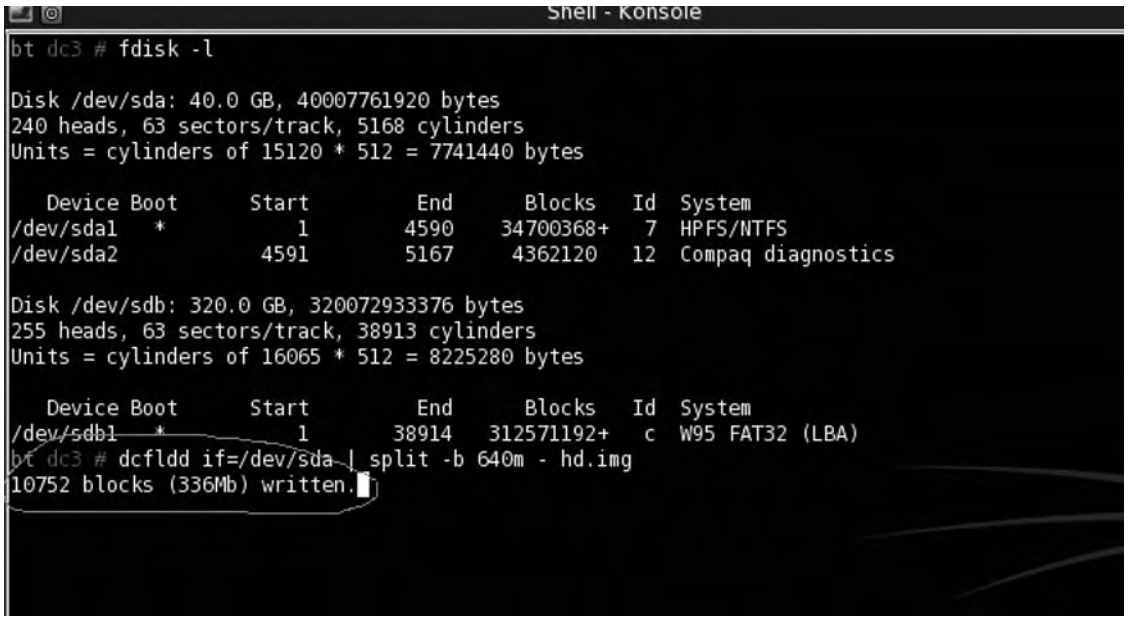
dcfldd was designed to be an open source computer forensic tool that would improve some of the shortcomings of the dd program. When the *hashwindow=0* option is specified, the dcfldd program will calculate the MD5 hash while the data is being copied, as shown in Figure 11.10. Using the *dcfldd* command with the *hashwindow* option will eliminate the extra step of using the md5sum program after making a bitstream copy. This can save a lot of time because hashing a drive with the *md5sum* command can take a long time.

Figure 11.10 The dcfldd Command with hashwindow=0 Switch Creating the MD5 Hash


Shell - Konsole
bt dc3 # fdisk -l
Disk /dev/sda: 40.0 GB, 40007761920 bytes
240 heads, 63 sectors/track, 5168 cylinders
Units = cylinders of 15120 * 512 = 7741440 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 4590 34700368+ 7 HPFS/NTFS
/dev/sda2 4591 5167 4362120 12 Compaq diagnostics
Disk /dev/sdb: 320.0 GB, 320072933376 bytes
255 heads, 63 sectors/track, 38913 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sdb1 * 1 38914 312571192+ c W95 FAT32 (LBA)
bt dc3 # dcfldd if=/dev/sda2 conv=noerror,sync hashwindow=0 | split -b 640m - hdd.img
136192 blocks (4256Mb) written.Total (md5): b328cb64dd19f4e0de29acf7c9c28cc8
136316+1 records in
136317+0 records out
bt dc3 #

Another feature of the dcfldd program is the status bar, highlighted in Figure 11.11. The status bar is important because it indicates how long the image processing will take. This is important to people who do a lot of imaging because the dd program does not give any indication of when the imaging process will be complete. Those of you who are comfortable with using imaging programs that have timers, such as ghost or EnCase, will appreciate the status bar of dcfldd. Like dd images, you can open dcfldd images and examine them using Autopsy, or proprietary tools such as EnCase, FTK, or X-Ways Forensics.

Figure 11.11 The dcfldd Status Bar


Shell - Konsole
bt dc3 # fdisk -l
Disk /dev/sda: 40.0 GB, 40007761920 bytes
240 heads, 63 sectors/track, 5168 cylinders
Units = cylinders of 15120 * 512 = 7741440 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 4590 34700368+ 7 HPFS/NTFS
/dev/sda2 4591 5167 4362120 12 Compaq diagnostics
Disk /dev/sdb: 320.0 GB, 320072933376 bytes
255 heads, 63 sectors/track, 38913 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sdb1 * 1 38914 312571192+ c W95 FAT32 (LBA)
bt dc3 # dcfldd if=/dev/sda | split -b 640m - hd.img
10752 blocks (336Mb) written.

Configuring & Implementing…

Imaging a Hard Drive with dcfldd

1. Turn the computer on and immediately enter the BIOS.
2. Verify that the CD/DVD-ROM drive is the first device listed in the boot order.
3. Insert the BackTrack CD and boot to BackTrack with the *bt nohd* command. Log in as root with a password of toor, and type **startx** to bring up the GUI.
4. Open a terminal session.
5. Type **fdisk –l**. You will see the drive letter and numbering designation of the source disk. (This is most likely HDA1 for a single IDE drive with one partition.)

Continued

6. Mount the drive as read-only:

```
ex:
cd /mnt
mkdir hda1
mount –o ro /dev/hda1 /mnt/hda1
```

7. Plug in the USB mass storage device.
8. Type **fdisk –l** again. You should see a new, added drive letter and partition, most likely SDA1. This is your destination disk. This drive needs to be mounted as read/write:

```
ex:
cd /mnt
mkdir sda1
mount –o rw /dev/sda1 /mnt/sda1
```

9. Navigate to a directory you created on the destination hard disk:

```
ex:
cd /mnt/sda1
mkdir evidence
cd evidence
```

10. **dcfldd=if/dev/hda bs=512 conv=sync,noerror hashwindow=0 | split –b 640m – hd.img**

dd_rescue

dd_rescue is a program designed by Kurt Garloff that will help you to recover data from a drive which has errors. The dd and dcfldd programs are great forensic tools that allow investigators to make bit-by-bit copies of nondamaged media. However, when media has input/output errors, many copying utilities such as dd, dcfldd, and Norton Symantec Ghost will error out. The *dd_rescue* command will continue to copy disks even when bad blocks are present. The dd_rescue program will not error out, unless you use the *–e* switch to specify the number of errors to stop. One of the nice features of dd_rescue is that it allows you to copy a drive from the end to the beginning.

One of the limitations of dd_rescue is that it can take a great deal of time to copy media with a large number of bad blocks. Even though the process can be time-consuming, dd_rescue has a more efficient way of dealing with errors. When you use dd with a block size greater than 512 bytes and the *conv=noerror,sync* switch, all bad blocks will be replaced with 0s on the destination media. In contrast, dd_rescue will copy data in 64k chunks until it reaches an error, at which point it will switch to copying data in 512-byte sectors.

After copying a specified number of error-free blocks, dd_rescue will switch back to copying data in 64k chunks again. dd_rescue's log file feature could be extremely useful if you need more information about the location of the bad sectors on the drive.

In general, dd_rescue will recover more data from damaged drives than the regular dd program. If a damaged drive may have critical evidence about an intrusion or workplace misconduct, you could use the dd_rescue tool to get as much information off that drive as possible. Copying drives with a large number of bad blocks is an extremely time-consuming process, so use your best discretion. You must weigh the amount of time you spend trying to recover data from disks with damaged blocks against the benefits of retrieving the information in the extra blocks. If you believe some of the data in the damaged blocks will "make or break the case," try to recover it. If you do not have the time or you do not have enough evidence, forget about the tedious process of copying a damaged drive. Additional tools are available that can assist you with recovering data from damaged disks. dd_rhelp is a tool you can add to a customized version of BackTrack that works well in conjunction with dd_rescue.

Configuring & Implementing...

Using dd_rescue to Save a Damaged Drive

1. Insert an additional hard drive into your system. If you can, get a drive that is from the same manufacturer as the original drive. The drive geometry should be the same. (If you cannot get the exact model, find a destination drive that is equal to or greater than the source disk in size.) Make sure the cables are jumpered properly and the new drive is recognized in the BIOS. Note the IDE cable and the jumper position of the source and destination drives.
2. Insert the BackTrack CD and boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
3. Use the information in Tables 11.1, 11.2, 11.3, and 11.4 to determine your source and destination lettering assignments.
4. Open dd_rescue from **K | BackTrack | Digital Forensics | Image Acquiring | DD_Rescue** to see the syntax and possible switches. Here is an example of how I would use dd_rescue to copy a damaged primary master drive to a secondary master (your syntax may vary, depending on the source and destination drives):

   ```
   ex: dd_rescue -A -v /dev/hda /dev/hdc
   ```

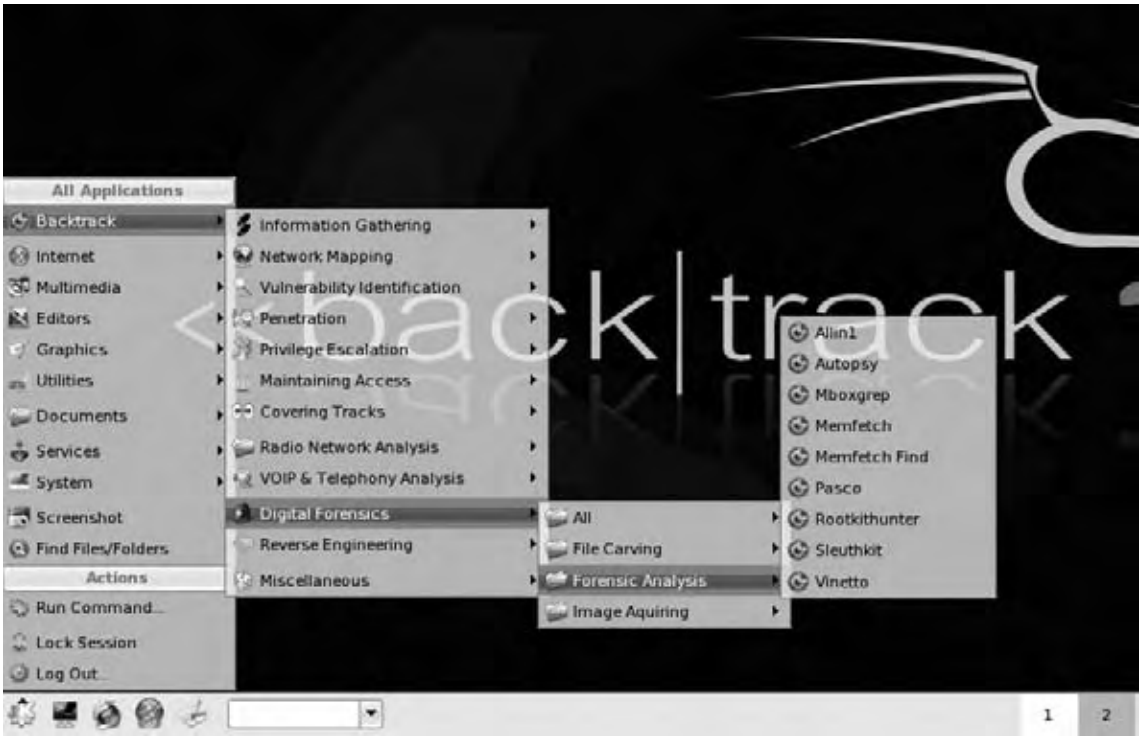
   You can add the *–l* switch to create a log file that will give information about the errors on the target disk.

Forensic Analysis

Forensic analysis can be useful for penetration testers to examine artifacts left by hackers, spyware, and viruses. Examining RAM, page or swap files, and deleted entries can be part of the forensic analysis process. These artifacts can provide a wealth of information about what hackers or intruders were doing to a system and how they may have tried to cover their tracks. As a penetration tester, it is necessary to be able to conduct forensic analysis to investigate and stop a current threat, and to prevent future ones. You cannot underestimate the number of individuals and organizations conducting these types of penetrations on various systems around the world. When you discover a hacked system, it is critical that you use forensic analysis tools to effectively examine what has happened and what you can do to fix the problem.

The BackTrack Security Distribution contains nine tools for forensic analysis; see Figure 11.12 for a display of those tools. Forensic analysis of media can be extremely important if an intrusion has occurred. When companies have security holes and their systems have been penetrated, forensic analysis of systems will help to give clues of what has been done to the systems and what can be done to prevent further attacks. Analysis of volatile and nonvolatile storage can give us clues regarding how and why hackers were able to penetrate a system, and what steps we can take to prevent them from doing more damage.

Figure 11.12 The Nine Tools Included in BackTrack for Forensic Analysis



Autopsy

Some of the most widely used forensic software packages include EnCase, FTK, and X-Ways Forensics. These proprietary packages are widely respected by law enforcement and by those in the forensic community; however, they come with large price tags. Autopsy is an open source forensic tool that has many of the same features as its commercial counterparts. Autopsy is also respected in both the forensic and law enforcement communities as a valuable tool for conducting forensic investigations. Brian Carrier is responsible for most of the development of The Sleuth Kit and Autopsy Browser. The Autopsy Browser is merely a GUI front end for The Sleuth Kit, which at one point was a commercial product from @stake, since acquired by Symantec.

The Autopsy Browser will allow you to open and examine both dd and dcfldd images as well as images that were created with EnCase, X-Ways Forensics, and FTK. When you select the Autopsy program from the BackTrack menu, a terminal window will pop up and direct you to paste http://localhost:9999/autopsy into a browser.

You open the Autopsy program by pasting the specified URL into either Konqueror or the Firefox browser. You need to open a case, give the case a name, and list the investigator or investigators. Select the **Add Host** button and fill in the Add a New Host information fields that you feel are pertinent to your investigation. Available fields include **Host Name**, **Description**, **Time Zone**, **Timeskew Adjustment**, and **Hash Paths**. You can add images by clicking **Add Image**, and then clicking **Add Image File**. The path to the image file will vary; if you completed the configuring and implementing exercise on dd or dcfldd earlier in this chapter and put it on your removable USB drive, it will most likely be /mnt/sda1/BackTrack/hd.img*. The asterisk is used when the image is split over several files.

Configuring & Implementing…

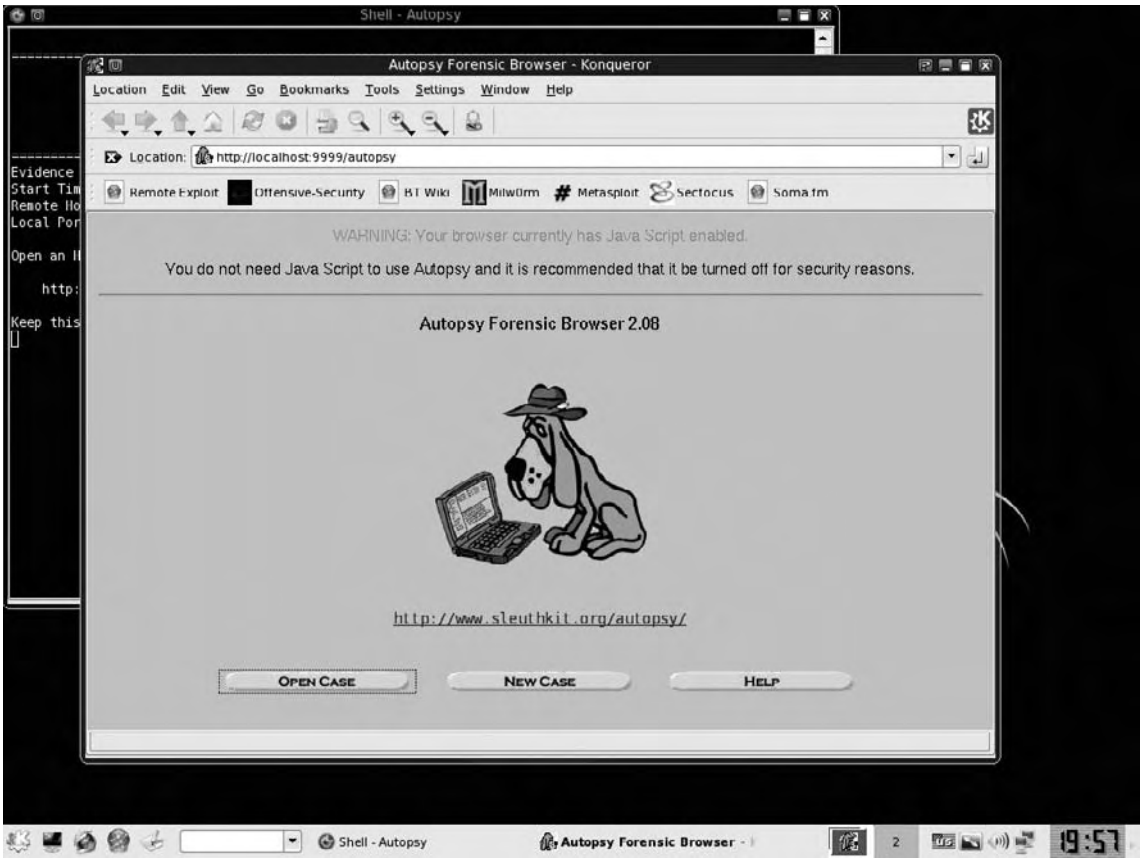
Adding a dd (or dcfldd) Image into Autopsy

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI. Connect your USB (or FireWire) drive with the dd and/or dcfldd image(s).
2. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Autopsy**.

Continued

3. Open the Konqueror or Firefox browser and go to the following URL: http://localhost:9999/autopsy. The screen will display Autopsy Forensic Browser 2.0, as shown in Figure 11.13.
4. Click the **Open Case** radio button. If cases are present on the system, they will be listed.
5. Click the **New Case** radio button.
6. Give your case a name and description, add yourself as an investigator, and click the **New Case** radio button again.
7. Click the **Add Host** radio button and fill in the fields that you feel are pertinent to your investigation. Available fields include **Host Name**, **Description**, **Time Zone**, **Timeskew Adjustment**, **Path of Alert Hash Database**, and **Path of Ignore Hash Database**. Then click the **Add Host** radio button at the bottom of the screen.
8. Click **Add Image**.
9. Click **Add Image File**.
   Under **Location:**, if you completed the configuring and implementing exercise on dd or dcfldd earlier in this chapter and put it on your removable USB drive, it will most likely be /mnt/sda1/BackTrack/hd.img*.
10. Select **Disk or Partition**. Under the Default import method, accept the default of **Symlink** and click **Next**.
11. At the Split Image Confirmation screen, verify that all parts of the image file are present and click **Next**.
12. Select **Calculate the hash value for this image** if you are interested in verifying file integrity. Under partition 1, note the mount point and file system type and click **Add**.
13. Click **OK** at the Testing Partitions screen.
14. At the **Select a volume to analyze** screen, click on **C:/** and select the **Analyze** radio button.
15. Click the **File Analysis** radio button and you will be in file browsing mode. By default, the files on the root directory will be displayed, along with a wealth of information about the files, including times written, accessed, changed, and metadata.

Figure 11.13 The Autopsy Forensic Browser 2.0



After you have added the image and selected the File Analysis radio button, all of the files on the root of the drive will be displayed. The default path is the root; you can enter the path to a folder, such as windows/system32 in the Directory Browser, to locate a folder. (Note that we're using the forward slash to stay consistent with Linux naming conventions, even though the backslash is consistent with traditional Windows naming conventions.) You also can use a search feature to search for certain types of files. For example, you can search

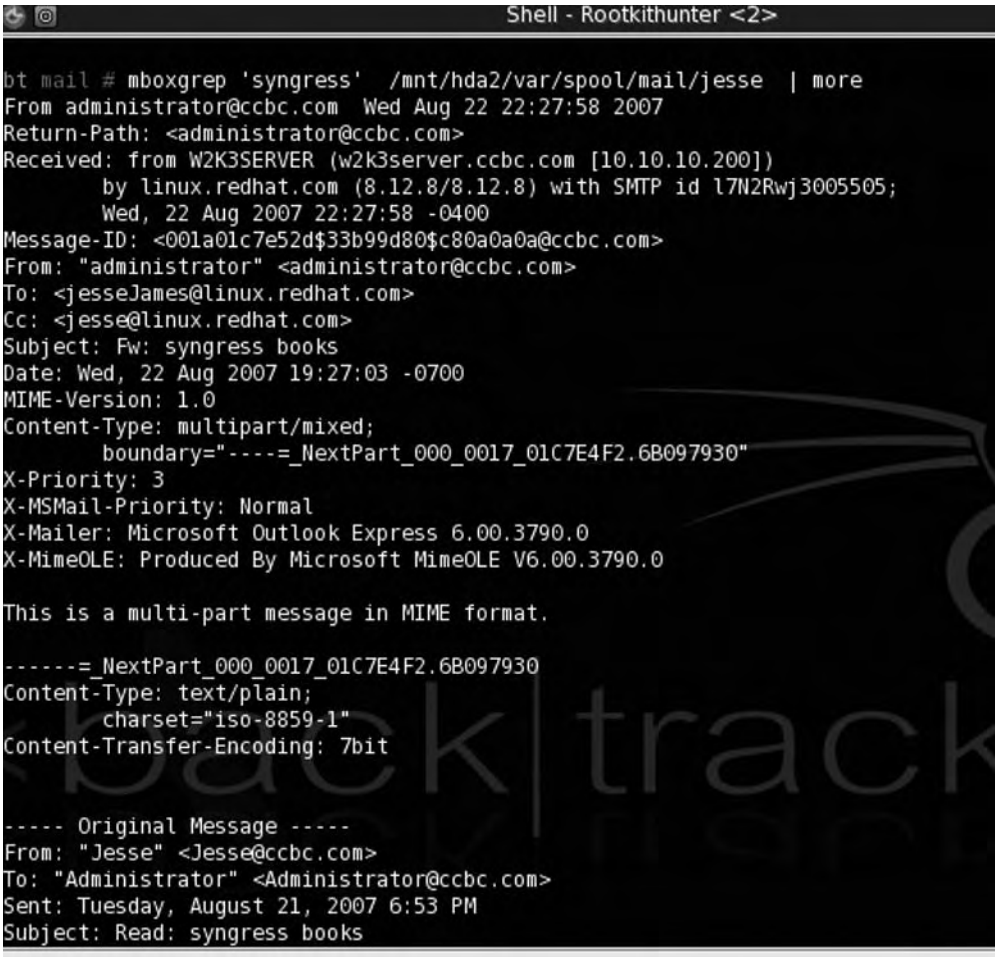
for all JPEG files with the .jpg extension by typing **\.jpg** in the File Name Search box and selecting the **Search** radio button. You can use the All Deleted Files radio button to find deleted files and to recover their contents. The Expand Directories button, located at the bottom left-hand pane, will give a tree view of all of the folders and their corresponding subfolders.

If you want to find an item within a file, such as a credit card or Social Security number, you need to conduct a keyword search. By default, the keyword search looks for both ASCII and Unicode strings. Keyword searches are case-sensitive by default; check the **Case Insensitive** box if you are unsure of the case of your expression. You can also perform a GREP of a regular expression in the keyword search, and a cheat sheet is provided if your GREP syntax has become rusty. (There are limitations to the keyword search capability, and they are described the GREP Search Limitations help file.) You may also find the predefined search terms included in the keyword search function to be useful. The five predefined keyword searches that you can perform are CC, IP, SSN1, SSN2, and Date.

mboxgrep

mboxgrep is a tool for use on Linux systems only; it was not designed to operate on Windows systems. It will allow you to query four types of Linux/UNIX mailbox folders: MBOX, MH, NNMH or NNML (GNUS), and MAILDIR (Q-MAIL). You also can look through the e-mail messages in these mailboxes. The *mboxgrep* command understands POSIX expressions. A person looking to retrieve e-mail messages can boot to the BackTrack Live CD on a system running Linux/UNIX with MBOX or other supported folder types. Then, by utilizing the mboxgrep program, the user can find e-mails from a particular domain or user, or with certain subject fields (see Figure 11.14). The user also can redirect the mboxgrep output to a text file and save it to a removable thumb drive or a floppy disk. Remember, however, that mboxgrep will be useful to you only if the system you are checking is using Linux/UNIX and one of the four supported types of e-mail folders.

Figure 11.14 Using the mboxgrep Command to Read Messages from a Linux Mailbox



Configuring & Implementing…

Using mboxgrep on a Linux System with MBOX Folders

1. Boot to BackTrack on your Linux/UNIX system with one of these four types of mailbox folders: MBOX, MH, NNMH or NNML (GNUS), or MAILDIR (Q-MAIL). Log in as root with a password of toor, and type **startx** to bring up the GUI.

Continued

2. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Mboxgrep**. Switch to the partition where your mailbox folders are located on the hard drive:

   ```
   ex: cd /mnt/hda2/var/spool/mail
   ```

3. Enter the following to find all the messages with the word "aol" in the mailbox of jessejames:

   ```
   mboxgrep 'aol' jessejames or mboxgrep 'aol' /mnt/hda2/var/spool/
   mail/jessejames
   ```

4. Enter the following to find all the messages with the word "mason" in the mailbox of jesse:

   ```
   mboxgrep 'mason' jesse or mboxgrep 'mason' /mnt/hda2/var/spool/mail/
   jesse
   ```

5. Enter the following to find all the messages without the word "gmail" in the mailbox of kim:

   ```
   mboxgrep -v 'gmail' kim or mboxgrep –v 'gmail' /mnt/hda2/var/spool/
   mail/kim
   ```

memfetch

Some of the most malicious programs hide themselves in RAM. A good penetration tester needs to be aware of the processes that are occurring in RAM to effectively analyze problems that may be occurring on a system. memfetch will allow you to capture the RAM from a process without affecting the process itself. Once you have captured and dumped the memory related to a certain process to files, you can further analyze what is happening on that system in volatile memory.

Booting to the BackTrack Live CD and using memfetch on a system would not be beneficial. When you boot to a Live CD, none of the normal operating system's services or processes are running. A major benefit of booting to a Live CD is that normal processes and services are not running, so fixing things and pinpointing issues can be less complicated. (This is somewhat comparable to safe mode in the world of Microsoft Windows.) However, if the operating system is infected with malware and the computer is booted to a Live CD environment, the malware processes tied to the normal operating system will most likely be inactive. So, to be able to fully analyze what is going on in a system's volatile memory, we need to boot the machine to its native operating environment. After the system is booted to the operating system and running in its normal mode of capacity, we will extract the memfetch utility from the BackTrack CD. Once we have copied the utility to the bin directory, we will use it to capture RAM from various processes.

Tools & Traps…

Memory Dumps from a Live System, Not a Live CD

Although it might seem like a good idea to boot to the BackTrack Live CD and run memfetch, it will not help to solve problems on the actual system. To figure out what is happening in volatile memory, you should boot to your normal operating system environment, install memfetch, and then do your memory dumps.

Configuring & Implementing…

Using memfetch on a Live Linux/UNIX System

1. Boot to your Linux/UNIX operating system and log in as root.
2. Insert your BackTrack CD-ROM.
3. Mount the CD-ROM if necessary; for example, **mount /mnt/cdrom**.
4. Go to the CD-ROM directory (*cd /mnt/cdrom*) and then to the BT/tools directory (*cd /BT/tools*; yes, the command is case-sensitive). Enter **ls** and verify that you see the lzm2dir file.
5. **./lzm2dir ../base/usr.lzm /root/**
6. Enter the following and verify that you see the memfetch file:

```
cd /root/usr/local/bin
ls memfetch
```

7. Enter the following:

```
cp memfetch /bin
```

   Keep it in the user binary executables for now; you can delete it later if you do not want it around.

8. Enter **ps aux** to identify a list of known processes on the system, or **top** (to identify a running list). Find the process ID number for which you want to do a memory dump.

Continued

9. mkdir memcapture1 (there may be a lot of files, and this will keep it neat) cd memcapture1
10. memfetch PID (will do a memory dump of the program with that process ID) ex: memfetch 7835

    Figure 11.15 shows how you will receive a series of memory dump files that will have a naming convention of map-xxx.bin and mem-xxx.bin. The mfetch.lst file will be a table of contents that explains which areas of memory map to the created files.
11. You can open the files with a hex editor and analyze them (I used Khexedit, shown in Figure 11.16), or you can use the Memfetch Find tool to look through all the files quickly.
12. For future memory captures, I suggest you create a new directory:

```
cd
mkdir memcapture2
cd memcapture2
```

Figure 11.15 Memory Dump Files Created with memfetch

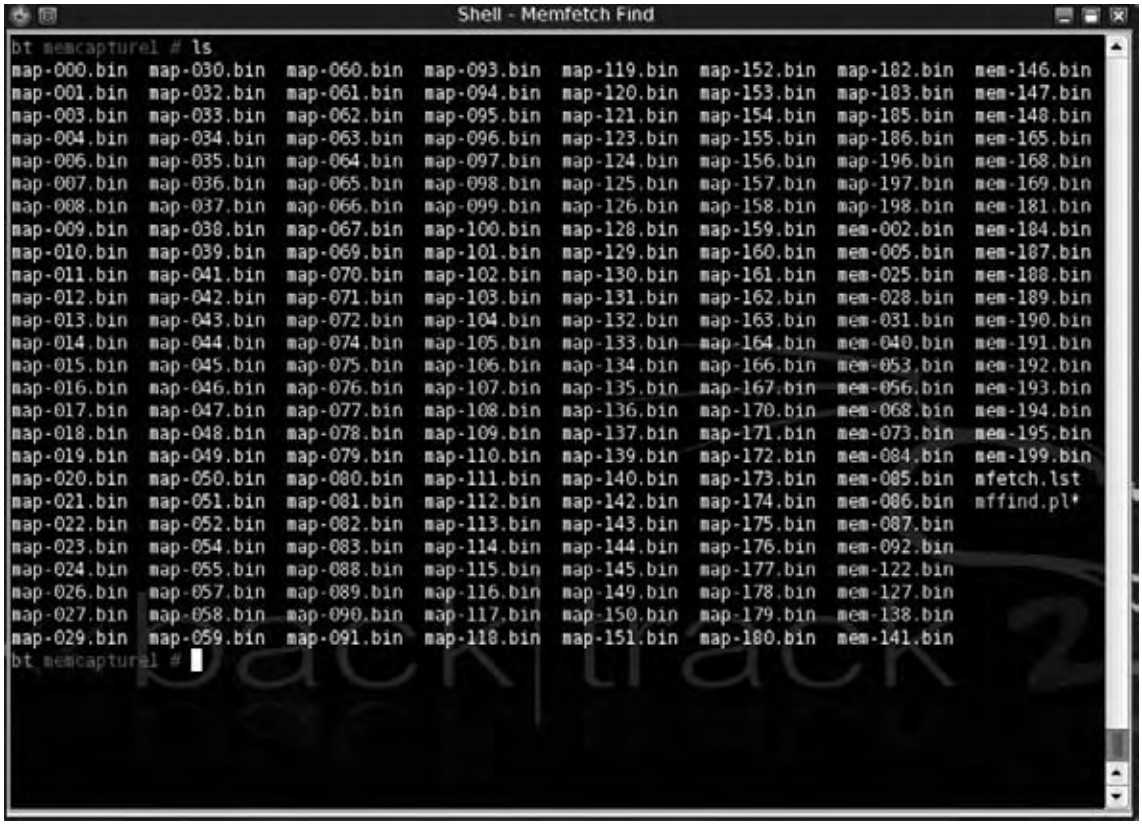
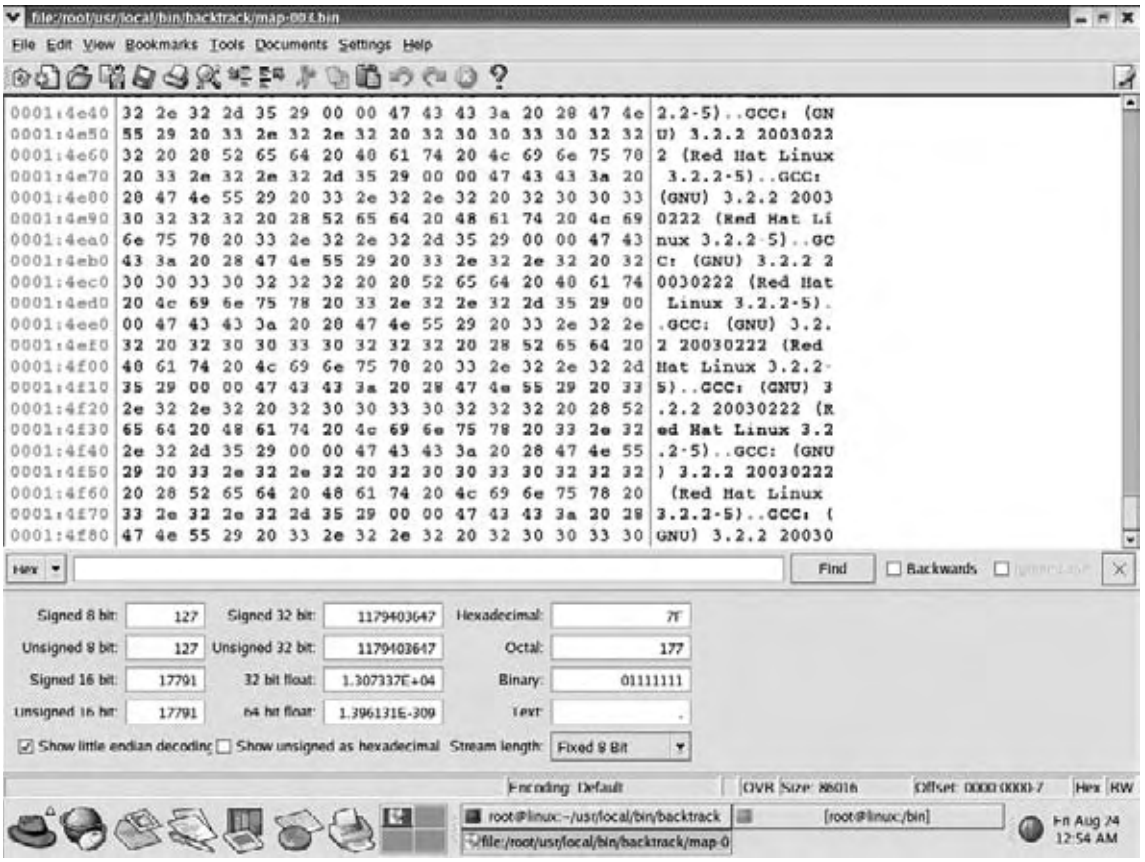

Shell - Memfetch Find
bt memcapture1 # ls
map-000.bin map-030.bin map-060.bin map-093.bin map-119.bin map-152.bin map-182.bin mem-146.bin
map-001.bin map-032.bin map-061.bin map-094.bin map-120.bin map-153.bin map-183.bin mem-147.bin
map-003.bin map-033.bin map-062.bin map-095.bin map-121.bin map-154.bin map-185.bin mem-148.bin
map-004.bin map-034.bin map-063.bin map-096.bin map-123.bin map-155.bin map-186.bin mem-165.bin
map-006.bin map-035.bin map-064.bin map-097.bin map-124.bin map-156.bin map-196.bin mem-168.bin
map-007.bin map-036.bin map-065.bin map-098.bin map-125.bin map-157.bin map-197.bin mem-169.bin
map-008.bin map-037.bin map-066.bin map-099.bin map-126.bin map-158.bin map-198.bin mem-181.bin
map-009.bin map-038.bin map-067.bin map-100.bin map-128.bin map-159.bin mem-002.bin mem-184.bin
map-010.bin map-039.bin map-069.bin map-101.bin map-129.bin map-160.bin mem-005.bin mem-187.bin
map-011.bin map-041.bin map-070.bin map-102.bin map-130.bin map-161.bin mem-025.bin mem-188.bin
map-012.bin map-042.bin map-071.bin map-103.bin map-131.bin map-162.bin mem-028.bin mem-189.bin
map-013.bin map-043.bin map-072.bin map-104.bin map-132.bin map-163.bin mem-031.bin mem-190.bin
map-014.bin map-044.bin map-074.bin map-105.bin map-133.bin map-164.bin mem-040.bin mem-191.bin
map-015.bin map-045.bin map-075.bin map-106.bin map-134.bin map-166.bin mem-053.bin mem-192.bin
map-016.bin map-046.bin map-076.bin map-107.bin map-135.bin map-167.bin mem-056.bin mem-193.bin
map-017.bin map-047.bin map-077.bin map-108.bin map-136.bin map-170.bin mem-068.bin mem-194.bin
map-018.bin map-048.bin map-078.bin map-109.bin map-137.bin map-171.bin mem-073.bin mem-195.bin
map-019.bin map-049.bin map-079.bin map-110.bin map-139.bin map-172.bin mem-084.bin mem-199.bin
map-020.bin map-050.bin map-080.bin map-111.bin map-140.bin map-173.bin mem-085.bin mfetch.lst
map-021.bin map-051.bin map-081.bin map-112.bin map-142.bin map-174.bin mem-086.bin mffind.pl*
map-022.bin map-052.bin map-082.bin map-113.bin map-143.bin map-175.bin mem-087.bin
map-023.bin map-054.bin map-083.bin map-114.bin map-144.bin map-176.bin mem-092.bin
map-024.bin map-055.bin map-088.bin map-115.bin map-145.bin map-177.bin mem-122.bin
map-026.bin map-057.bin map-089.bin map-116.bin map-149.bin map-178.bin mem-127.bin
map-027.bin map-058.bin map-090.bin map-117.bin map-150.bin map-179.bin mem-138.bin
map-029.bin map-059.bin map-091.bin map-118.bin map-151.bin map-180.bin mem-141.bin
bt memcapture1 #

Figure 11.16 Opening a Memory Dump File Created with memfetch in Khexedit



Memfetch Find

Memfetch Find, or mffind.pl, is a program that will allow you to search through memfetch dump files using Perl expressions. You can use this useful Perl script to find the memory locations of certain expressions. The code in the memfetch file was contributed by spoonm@ghettohackers.net. Figure 11.17 shows the code written to the mffind.pl file. Before using the file, give the root user execute permissions on the file, as shown in Figure 11.18.

Figure 11.17 The Code of the Memfetch Find mffind.pl File

```
die "$0 <regex>\n" if(!@ARGV);
undef $/;
my @files;
open(LISTFILE, '<mfetch.lst') or die "Cannot open mfetch.lst: $!\n";
my $listFile = <LISTFILE>;
while($listFile =~ /\]\s+(.*?):.*?range (.*?) /sg) {
  push(@files, [ $1, $2 ]);
}
close(LISTFILE);

foreach (@files) {
  open(INFILE, "<$_->[0]") or die "Cannot open $_->[0]: $!\n";
  my $file = <INFILE>;
  while($file =~ /($ARGV[0])/sg) {
    my $location = pos($file) - length($1);
    print "[$1] $location + $_->[1] -> ",
      sprintf("%#08x\n", $_->[1] + $location);
  }
  close(INFILE);
}
```

Configuring & Implementing…

Using Memfetch Find on a memfetch Memory Dump

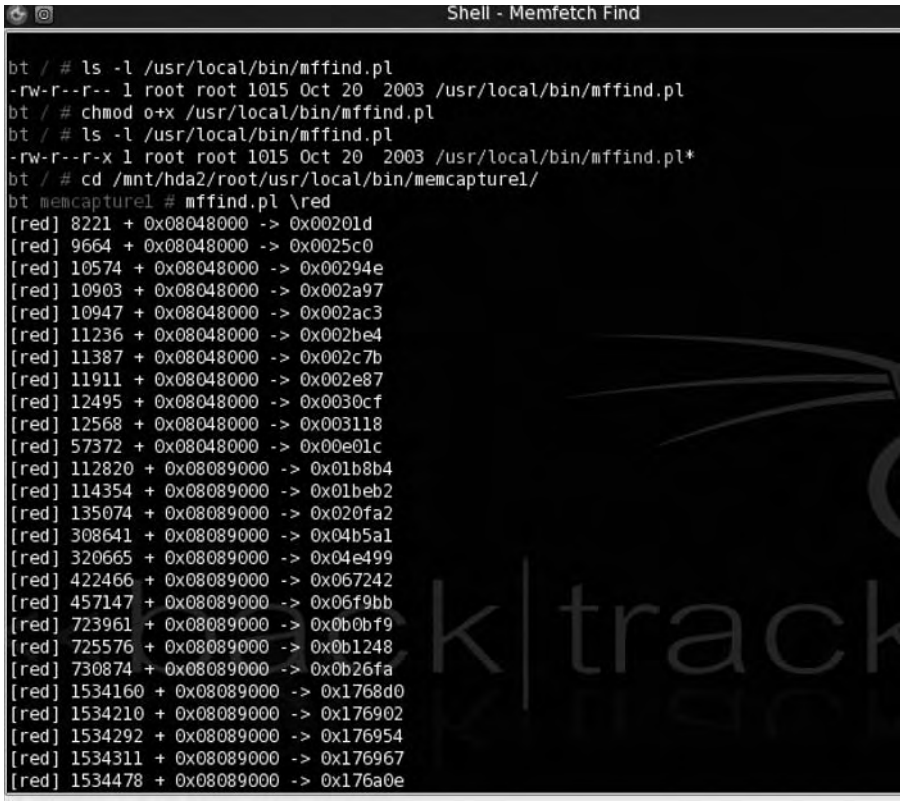
1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
2. Enter the following and verify that the user (root) has execute permissions:

   ```
   ls –l /usr/local/bin/mffind.pl
   chmod u+x /usr/local/bin/mffind.pl
   ls –l /usr/local/bin/mffind.pl
   ```

3. Go into the directory of the memory dump files (*cd /root/usr/local/bin/memcapture1*).
4. mffind.pl\ expression or mffind.pl\ expression+otherexpression you are looking for.

   For example, *mffind.pl\ red* searches through the memory dump files for the word "red"; *mffind.pl\ hat* searches through the memory dump files for the word "hat"; and *mffind.pl\ red+hat* searches through the memory dump files for the words "red hat".

Figure 11.18 Using Memfetch Find



pasco

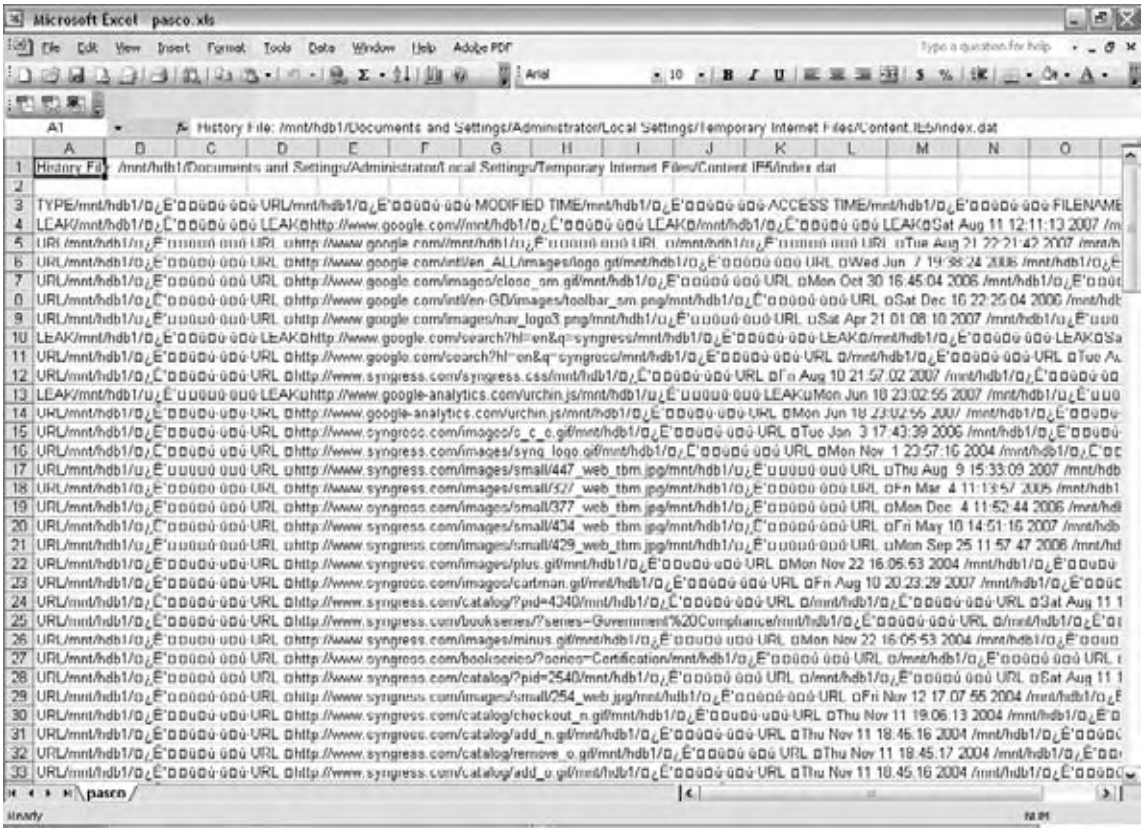
The pasco utility included with BackTrack is a forensic tool for reading the index.dat files that are created by Microsoft's Internet Explorer (pasco comes from the Latin word *browse*). Other Internet browser software, such as Firefox and Safari, do not use index.dat files. The index.dat file leaves traces of a user's browsing history. Browsing history can often indicate whether criminal activity has taken place, or it can help to determine what tasks users with access to a system were trying to accomplish. The index.dat files are located in several areas on a hard drive. Some index.dat files are associated with each user's profile, and these files are located throughout a user's home folder as well as other areas on the drive.

The *find* command will allow you to locate all of the index.dat files on a target drive. Once you've located the files, you can use the pasco utility to extract Internet Explorer browser information from the files. pasco will organize the browsing session into the following seven categories: type, URL, modified time, access time, filename, directory, and HTTP header.

Using the *–d* option with pasco—also known as undeletion mode—will often uncover even more browsing history information. In addition, you can use the *–t* option to specify

the comma as the delimiter between the seven fields. You can also redirect pasco output to a file and open that file in BackTrack or save it to removable media and open it later in Notepad or Microsoft Excel, as displayed in Figure 11.19.

Figure 11.19 pasco Output Redirected into a File and Opened with Microsoft Excel


Microsoft Excel - pasco.xls
File Edit View Insert Format Tools Data Window Help Adobe PDF
History File: /mnt/hdb1/Documents and Settings/Administrator/Local Settings/Temporary Internet Files/Content.IE5/index.dat
History File: /mnt/hdb1/Documents and Settings/Administrator/Local Settings/Temporary Internet Files/Content.IE5/index.dat
TYPE/mnt/hdb1/ MODIFIED TIME/mnt/hdb1/ ACCESS TIME/mnt/hdb1/ FILENAME
pasco
Ready

Configuring & Implementing…

Using pasco to Read index.dat Files

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
2. find / -name index.dat

Continued

3. pasco full path to the file in quotes:

ex: pasco "/mnt/hda1/Documents and Settings/Administrator/UserData/index.dat"

4. To separate the output into a comma-delimited format enter the following:

```
pasco -t, "/mnt/hda1/Documents and Settings/Administrator/UserData/
index.dat"
```

5. Try to extract even more browsing information by entering the following:

```
pasco -d "/mnt/hda1/Documents and Settings/Administrator/UserData/
index.dat"
```

6. Redirect output into a text file by entering the following:

```
pasco "/mnt/hda1/Documents and Settings/Administrator/UserData/
index.dat" > brower.txt
```

Rootkit Hunter

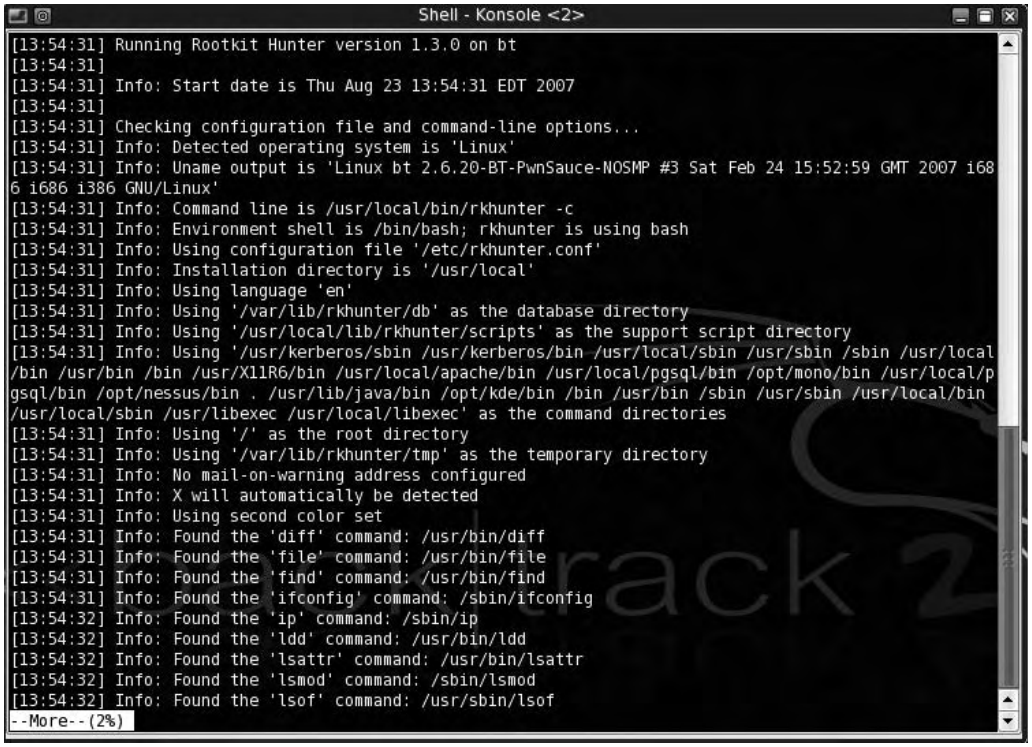
A colleague of mine once told me about a case in which a production machine running Windows 2000 Server kept blue-screening. Eventually, after a lot of hard work and time, the problem turned out to be a rootkit installed by individuals who had penetrated the system. Rootkits are tools that individuals install for malicious purposes in an attempt to control a system and conceal their presence. If a penetrator installs a rootkit correctly, normal users (and even some advanced users) will not be able to detect his presence. That is where the Rootkit Hunter tool in BackTrack can come in handy. Rootkit Hunter, or rkhunter, is a tool for detecting rootkits, backdoors, and sniffers. Rootkit Hunter was designed for use on Linux/UNIX-based systems only; you can also use it on Solaris and Macintosh systems running OS 10.3.4–10.3.8. It was not designed to operate on Windows systems. You can find a full list of supported platforms at www.rootkit.nl/projects/rootkit_hunter.html.

The version of Rootkit Hunter that is included with BackTrack, Rootkit Hunter 1.2.8, removes 58 rootkits including Knark and T0rn; for a complete list of rootkits that Rootkit Hunter removes, refer to the tool's Web site. The newest version of Rootkit Hunter, version 1.3.0, checks for 114 different rootkits. If you customize your edition of BackTrack, you may want to include the latest version of rkhunter. Rootkit Hunter uses a series of known good MD5 hashes to verify files, such as the user and administrative binary files. To use Rootkit Hunter, boot the supported system to the BackTrack CD, do a chroot to the local system (usually /mnt/hda2), and run *rkhunter -c*. You will be asked to

press **Enter** several times during the scanning process. Rootkit Hunter will relay a series of messages to you, including OK, Not Found, None Found, Found, Bad, and Warning. Warning, Bad, and Found do not necessarily indicate that a rootkit is present; they could indicate that a port is open.

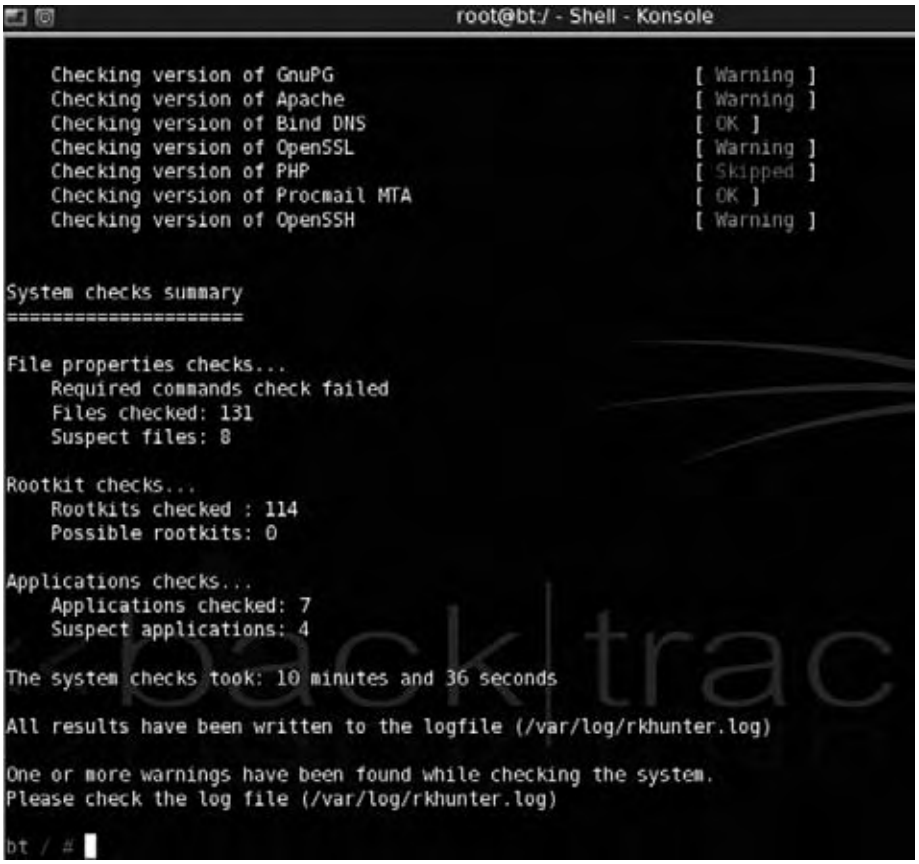
False positives are possible with rkhunter, so if you see something suspicious, it may warrant further research. You should decrease the number of false positives when you run rkhunter from the BackTrack environment. If you run rkhunter from the actual operating system, running applications will often trigger false positives. If you run version 1.2.8 with the *–createlogfile* switch, the log file /var/log/rkhunter.log, shown in Figure 11.20, will be able to give you more information about what rkhunter detected.

Figure 11.20 The /var/log/rkhunter.log Created by Rootkit Hunter


Shell - Konsole <2>
[13:54:31] Running Rootkit Hunter version 1.3.0 on bt
[13:54:31]
[13:54:31] Info: Start date is Thu Aug 23 13:54:31 EDT 2007
[13:54:31]
[13:54:31] Checking configuration file and command-line options...
[13:54:31] Info: Detected operating system is 'Linux'
[13:54:31] Info: Uname output is 'Linux bt 2.6.20-BT-PwnSauce-NOSMP #3 Sat Feb 24 15:52:59 GMT 2007 i68
6 i686 i386 GNU/Linux'
[13:54:31] Info: Command line is /usr/local/bin/rkhunter -c
[13:54:31] Info: Environment shell is /bin/bash; rkhunter is using bash
[13:54:31] Info: Using configuration file '/etc/rkhunter.conf'
[13:54:31] Info: Installation directory is '/usr/local'
[13:54:31] Info: Using language 'en'
[13:54:31] Info: Using '/var/lib/rkhunter/db' as the database directory
[13:54:31] Info: Using '/usr/local/lib/rkhunter/scripts' as the support script directory
[13:54:31] Info: Using '/usr/kerberos/sbin /usr/kerberos/bin /usr/local/sbin /usr/sbin /sbin /usr/local
/bin /usr/bin /bin /usr/X11R6/bin /usr/local/apache/bin /usr/local/pgsql/bin /opt/mono/bin /usr/local/p
gsql/bin /opt/nessus/bin . /usr/lib/java/bin /opt/kde/bin /bin /usr/bin /sbin /usr/sbin /usr/local/bin
/usr/local/sbin /usr/libexec /usr/local/libexec' as the command directories
[13:54:31] Info: Using '/' as the root directory
[13:54:31] Info: Using '/var/lib/rkhunter/tmp' as the temporary directory
[13:54:31] Info: No mail-on-warning address configured
[13:54:31] Info: X will automatically be detected
[13:54:31] Info: Using second color set
[13:54:31] Info: Found the 'diff' command: /usr/bin/diff
[13:54:31] Info: Found the 'file' command: /usr/bin/file
[13:54:31] Info: Found the 'find' command: /usr/bin/find
[13:54:31] Info: Found the 'ifconfig' command: /sbin/ifconfig
[13:54:32] Info: Found the 'ip' command: /sbin/ip
[13:54:32] Info: Found the 'ldd' command: /usr/bin/ldd
[13:54:32] Info: Found the 'lsattr' command: /usr/bin/lsattr
[13:54:32] Info: Found the 'lsmod' command: /sbin/lsmod
[13:54:32] Info: Found the 'lsof' command: /usr/sbin/lsof
--More--(2%)

As a security professional, it would be a good idea to use the latest version of Rootkit Hunter on your customized version of BackTrack. For one, it detects almost twice as many types of rootkits as version 1.2.8 does. The newest version also includes a convenient final report which indicates the possible rootkits, as shown in Figure 11.21. The newest version also automatically generates the /var/log/rkhunter.log file without any additional switches.

Figure 11.21 Checking for 114 Different Rootkits with Rootkit Hunter Version 1.3.0


root@bt:/ - Shell - Konsole
Checking version of GnuPG [Warning]
Checking version of Apache [Warning]
Checking version of Bind DNS [OK]
Checking version of OpenSSL [Warning]
Checking version of PHP [Skipped]
Checking version of Procmail MTA [OK]
Checking version of OpenSSH [Warning]
System checks summary
File properties checks...
Required commands check failed
Files checked: 131
Suspect files: 8
Rootkit checks...
Rootkits checked : 114
Possible rootkits: 0
Applications checks...
Applications checked: 7
Suspect applications: 4
The system checks took: 10 minutes and 36 seconds
All results have been written to the logfile (/var/log/rkhunter.log)
One or more warnings have been found while checking the system.
Please check the log file (/var/log/rkhunter.log)
bt / #

The Sleuth Kit

The Sleuth Kit is a group of 27 command-line tools, shown in Figure 11.22, which can help you to conduct computer forensic investigations. After you boot to the BackTrack Live CD environment, you will find these tools in the /usr/local/sleuthkit/bin directory. These tools will work on media with the following types of file systems: FAT12, FAT16, FAT32, NTFS, UFS1, UFS2, EXT2, EXT3, and ISO 9660. The Autopsy Browser, developed by Brian Carrier, is actually a front-end GUI for many of the Sleuth Kit tools. The Allin1 tool in BackTrack, developed by David Berger, is also a front end for some of the Sleuth Kit tools. The Sleuth Kit toolset will run on Linux, UNIX, Mac, and Solaris platforms. The Sleuth Kit tools are primarily used for analyzing disks and partitions, and are written to work with disks, partitions, files, or a combination of the three. An abundance of information about the

Sleuth Kit tools is available on the official Web site, www.sleuthkit.org/. Make sure you launch The Sleuth Kit from the BackTrack program menu, or open a terminal and navigate to the /usr/local/sleuthkit/bin directory. The Sleuth Kit binary files are not included as part of the user (bin) or system administrator (sbin) directory.

The *file* command is a useful Sleuth Kit tool that will probe a disk, partition, or file for information. Output from the *file* command, shown in Figure 11.23, can give you insight into what type of image they are using (sometimes you need the image type or details for the switches when a user inputs a command). Another useful tool for analyzing files is the *srch_strings* command. The *srch_strings* command will display strings in the file; it can also scan the whole file, not just the data portion. Two other tools you can use on disks, partitions, and files are md5 and sha1, both for retrieving hash values. The md5 tool works the same way as the *md5sum* command. Meanwhile, you can use the *sha1* command if you prefer a different hash or an additional one. The syntax is simply *md5 filename* or *sha1 filename*. Computer forensic investigators use hash values to verify that image files have the same data structure as the original suspect's media. This helps to convince juries that the investigator is working with an exact copy of the media and has not tampered with the original evidence.

Figure 11.22 The 27 Command-Line Forensic Tools of The Sleuth Kit

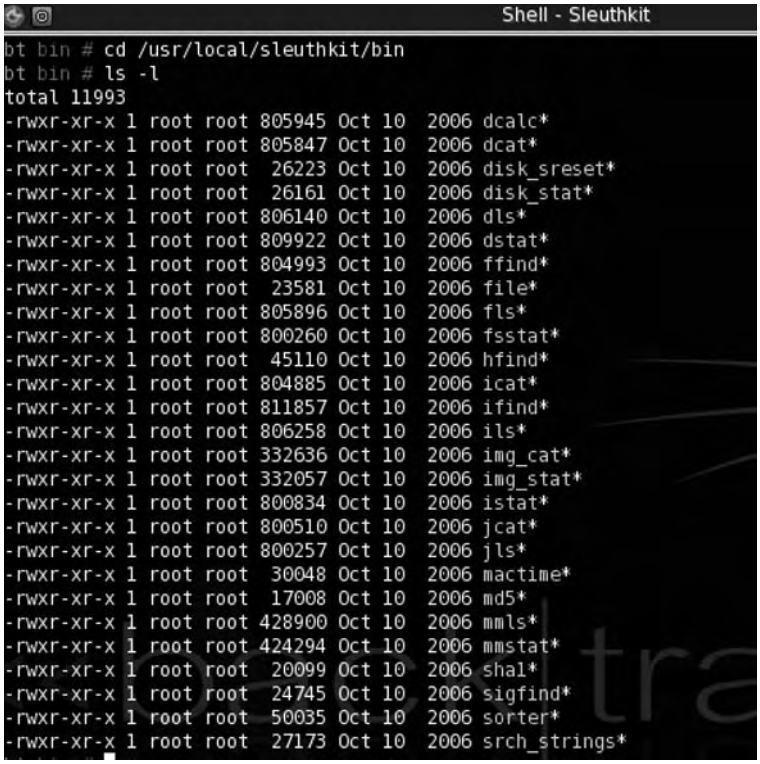
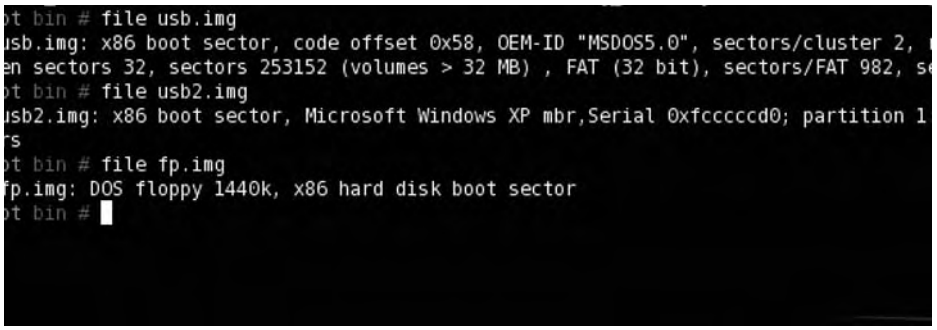

Shell - Sleuthkit
bt bin # cd /usr/local/sleuthkit/bin
bt bin # ls -l
total 11993
-rwxr-xr-x 1 root root 805945 Oct 10 2006 dcalc*
-rwxr-xr-x 1 root root 805847 Oct 10 2006 dcat*
-rwxr-xr-x 1 root root 26223 Oct 10 2006 disk_sreset*
-rwxr-xr-x 1 root root 26161 Oct 10 2006 disk_stat*
-rwxr-xr-x 1 root root 806140 Oct 10 2006 dls*
-rwxr-xr-x 1 root root 809922 Oct 10 2006 dstat*
-rwxr-xr-x 1 root root 804993 Oct 10 2006 ffind*
-rwxr-xr-x 1 root root 23581 Oct 10 2006 file*
-rwxr-xr-x 1 root root 805896 Oct 10 2006 fls*
-rwxr-xr-x 1 root root 800260 Oct 10 2006 fsstat*
-rwxr-xr-x 1 root root 45110 Oct 10 2006 hfind*
-rwxr-xr-x 1 root root 804885 Oct 10 2006 icat*
-rwxr-xr-x 1 root root 811857 Oct 10 2006 ifind*
-rwxr-xr-x 1 root root 806258 Oct 10 2006 ils*
-rwxr-xr-x 1 root root 332636 Oct 10 2006 img_cat*
-rwxr-xr-x 1 root root 332057 Oct 10 2006 img_stat*
-rwxr-xr-x 1 root root 800834 Oct 10 2006 istat*
-rwxr-xr-x 1 root root 800510 Oct 10 2006 jcat*
-rwxr-xr-x 1 root root 800257 Oct 10 2006 jls*
-rwxr-xr-x 1 root root 30048 Oct 10 2006 mactime*
-rwxr-xr-x 1 root root 17008 Oct 10 2006 md5*
-rwxr-xr-x 1 root root 428900 Oct 10 2006 mmls*
-rwxr-xr-x 1 root root 424294 Oct 10 2006 mmstat*
-rwxr-xr-x 1 root root 20099 Oct 10 2006 sha1*
-rwxr-xr-x 1 root root 24745 Oct 10 2006 sigfind*
-rwxr-xr-x 1 root root 50035 Oct 10 2006 sorter*
-rwxr-xr-x 1 root root 27173 Oct 10 2006 srch_strings*

Figure 11.23 The file Command Listing Pertinent Information about the Images


bt bin # file usb.img
usb.img: x86 boot sector, code offset 0x58, OEM-ID "MSDOS5.0", sectors/cluster 2,
en sectors 32, sectors 253152 (volumes > 32 MB) , FAT (32 bit), sectors/FAT 982,
bt bin # file usb2.img
usb2.img: x86 boot sector, Microsoft Windows XP mbr,Serial 0xfcccccd0; partition 1
rs
bt bin # file fp.img
fp.img: DOS floppy 1440k, x86 hard disk boot sector
bt bin #

The mmls tool, the output of which is shown in Figure 11.24, is useful when looking at a disk's partition table. This tool can analyze five partition tables, including DOS, BSD, Mac, Sun, and GUID. Unallocated spaces on the disk are also listed in the output of the mmls chart. If you do not know what types of partitions you are dealing with, you can use the mmstat tool to display information regarding the partition types.

Figure 11.24 The mmls Command Showing a Disk's Partition Table

```
DOS Partition Table
Offset Sector: 0
Units are in 512-byte sectors

     Slot    K         End          Length       Description
00:  -----   0000000000   0000000000   0000000001   Primary
Table (#0)
01:  -----   0000000001   0000000031   0000000031   Unallocated
02:  00:00   0000000032   0000253183   0000253152   Win95 FAT32
(0x0B)
```

Computer forensic investigators have to stay cognizant of Host Protected Areas (HPAs) on disks. The HPA is an area of the disk, hidden from the BIOS, that usually contains items such as vendor restoration utilities. Most investigators want to include this part of a drive in their image because a person with advanced knowledge of hard drives might use this area to hide data. Two Sleuth Kit tools deal specifically with the HPA of a disk: disk_sreset and disk_stat. The disk_sreset tool will temporarily remove the HPA from the disk (it will be added back upon reboot). The disk_stat tool will show whether an HPA exists on the disk.

The fsstat tool gives you a wealth of information about a file system image, as shown in Figure 11.25. When you use this tool, you will need to specify the file system and image

types. For a list of supported file system types and image types, type **fsstat –i list** (for images) and **fsstat –f list** (for file systems). Information included from the output of an *fsstat* command includes file system information, metadata information, content information, and FAT contents (if it is a FAT file system). The *fls* command will display all files and folders on a partition, including deleted files and folders. It is a useful tool for analyzing what entries are still present on a file system.

Figure 11.25 Information about the File System from the fsstat Command-Line Tool

```
bt bin # fsstat  -i raw fp.img
FILE SYSTEM INFORMATION
--------------------------------------------
File System Type: FAT12

OEM Name: MSWIN4.0
Volume ID: 0x18e63915
Volume Label (Boot Sector): BOOT98SE
Volume Label (Root Directory):
File System Type Label: FAT12

Sectors before file system: 0

File System Layout (in sectors)
Total Range: 0 - 2879
* Reserved: 0 - 0
** Boot Sector: 0
* FAT 0: 1 - 9
* FAT 1: 10 - 18
* Data Area: 19 - 2879
** Root Directory: 19 - 32
** Cluster Area: 33 - 2879

METADATA INFORMATION
--------------------------------------------
Range: 2 - 45554
Root Directory: 2

CONTENT INFORMATION
--------------------------------------------
Sector Size: 512
Cluster Size: 512
Total Cluster Range: 2 - 2848

FAT CONTENTS (in sectors)
--------------------------------------------
33-467 (435) -> EOF
468-468 (1) -> EOF
469-470 (2) -> EOF
471-654 (184) -> EOF
656-687 (32) -> EOF
688-764 (77) -> EOF
765-765 (1) -> EOF
766-1266 (501) -> EOF
1430-1446 (17) -> EOF
1447-1469 (23) -> EOF
```

Two tools that are helpful for determining information about an image are img_stat and img_cat. The img_stat tool with the *–t* switch will indicate whether the image is raw, aff, afd, afm, ewf, or split. Drives imaged with the dd and dcfldd tools will most likely be in the raw or split format. Once the image type is determined, you can use the *img_stat* command with *–i imagetype* to determine the size of the image file. The img_cat tool is useful if you want to display the raw data of the image file on the screen.

The *jcat* and *jls* Sleuth Kit commands work with journaling file systems. A journaling file system keeps a list of changes that are being made to the disk. If power is lost before the changes are made, the journal will pick up where it left off when the system comes back online. NTFS and EXT3 are good examples of journaling file systems. The jls tool lists the file system entries and the jcat tool will display a journal block.

The Sleuth Kit also includes the metadata tools icat, ils, ifind, ffind, and istat, which all reveal information about the metadata of files. You use the ifind and ffind tools to find the file where a string is located. You also can use these two tools within the Autopsy Browser, in conjunction with GREP, to locate the files that contain certain expressions. Tools from The Sleuth Kit that deal with data units include dstat, dls, dcat, and dcalc. Other tools include hfind, mactime, sorter, and sigfind. For more details on any of these additional tools, you can visit www.sleuthkit.org/sleuthkit/tools.php.

Configuring & Implementing…

Using the Sleuth Kit Command-Line Tools

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
2. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Sleuthkit**. Optional: It might be a good idea to make a disk image from a USB stick to do a practice run, preferably one that is smaller—say, 512-, 256-, or 128 MB with a FAT file system. Simply insert the USB stick, enter **fdisk –l** (should be SDA1 or similar), and then enter **dd if=/dev/sda of=usb.img**. This will put the image of the USB disk in the current directory.
3. md5 /dev/sda
   md5 usb.img
   The md5 values should match

Continued

4. sha1 /dev/sda
   sha1 usb.img
   The sha1 values should match
5. Get information about the raw image by using the *file* command:

   ```
   file usb.img
   ```

6. Get the partition table information from the disk:

   ```
   mmls -t dos -i raw usb.img
   ```

   Optional: Making a partition image from that same USB stick to do a practice run might be a good idea. Type **dd if=/dev/sda1 of=usb2.img** to put the image of the USB partition in the current directory.
7. md5 /dev/sda1
   md5 usb2.img
   The md5 values should match
8. sha1 /dev/sda1
   sha1 usb2.img
9. The SHA1 values should match. Get information about the raw image by using the *file* command:

   ```
   file usb2.img
   ```

10. Use *fsstat* to get a wealth of information about the file system image:

    ```
    fsstat -f fat -i raw usb.img
    ```

11. Use the *fls* command to display all files and folders, including deleted files, on a partition:

    ```
    fls -f fat -I raw usb2.img
    ```

The Sleuth Kit Continued: Allin1 for The Sleuth Kit

The Allin1 tool is a front end for The Sleuth Kit that will allow you to perform several Sleuth Kit procedures from one convenient location. Allin1 has options that allow you to perform up to 10 different tasks, as shown in Figure 11.26. It is important to note that BackTrack does not set the correct path to the Sleuth Kit binaries by default. The correct path should be /usr/local/sleuthkit/bin. Also, you need a configuration file, called host.aut. You can use a text editor such as vi or nano to construct one of these files, as demonstrated in Figure 11.27, or if you started a case using the Autopsy Browser, a configuration file will be constructed for you, which is recommended. After the file is constructed, show hidden

files in the Allin1 browser to find the file in its parent folder, .autopsy. Point to the host.aut file and click the **Info** button, and you'll see details regarding the image file. Click **OK**, and the program will freeze until it has completed its tasks, at which time it will indicate that it has finished. In the case folder, a log file called allin1.log will be created and recovered items will be extracted to the output folder. The configuration of the program may frustrate you initially, but the program is extremely beneficial once you've mastered it.

Figure 11.26 The Allin1 Tool for The Sleuth Kit

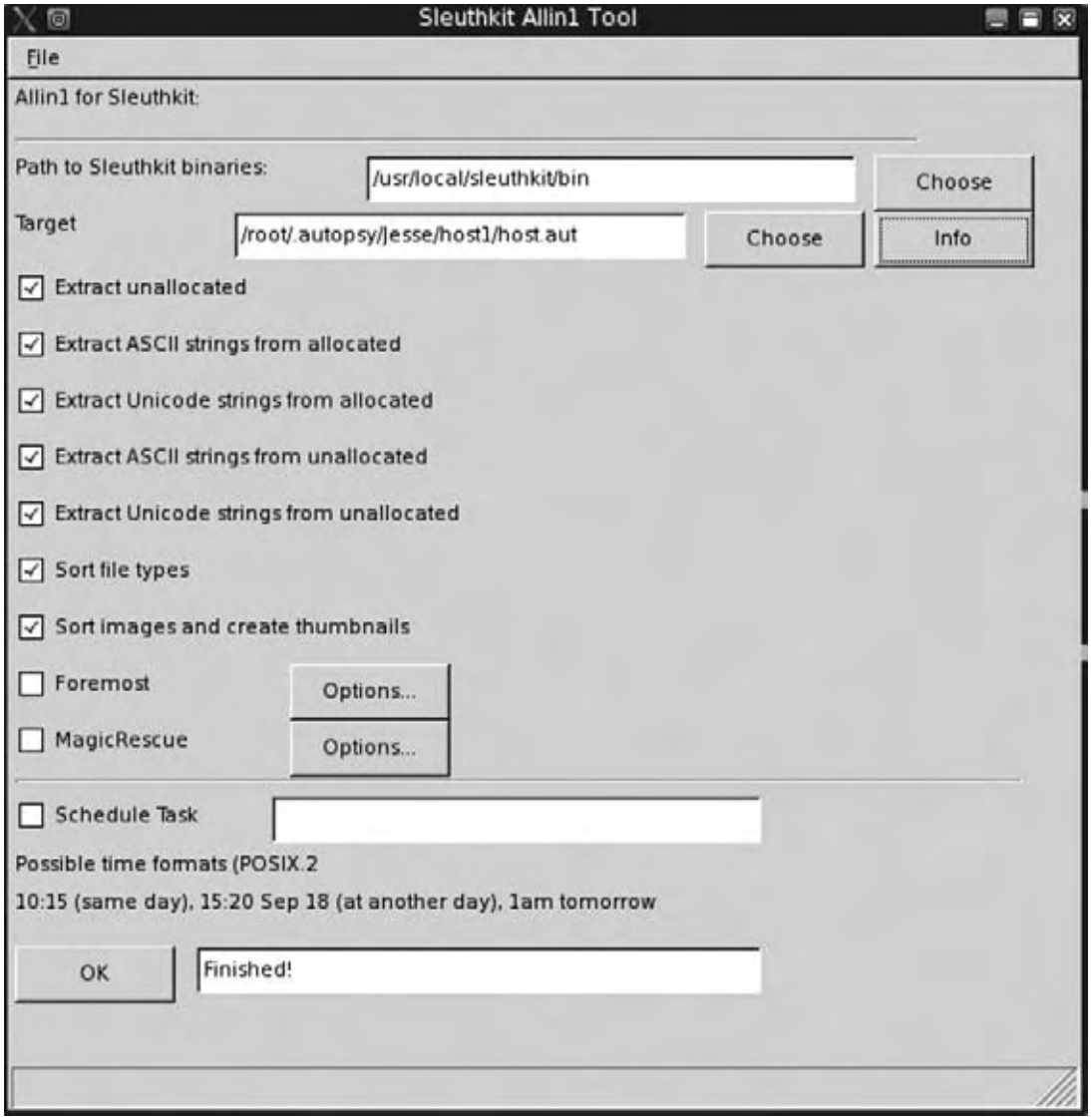


Figure 11.27 Sample Allin1 host.aut File

```
# Autopsy host config file
# Case: Syngress          Host: host1
# Created: Sun Aug 26 16:40:54 2007

timeskew  0
image  img1  raw  images/usb.img
disk  vol1  img1    dos
part  vol2  img1    32  253183    fat32  C:
```

Configuring & Implementing…

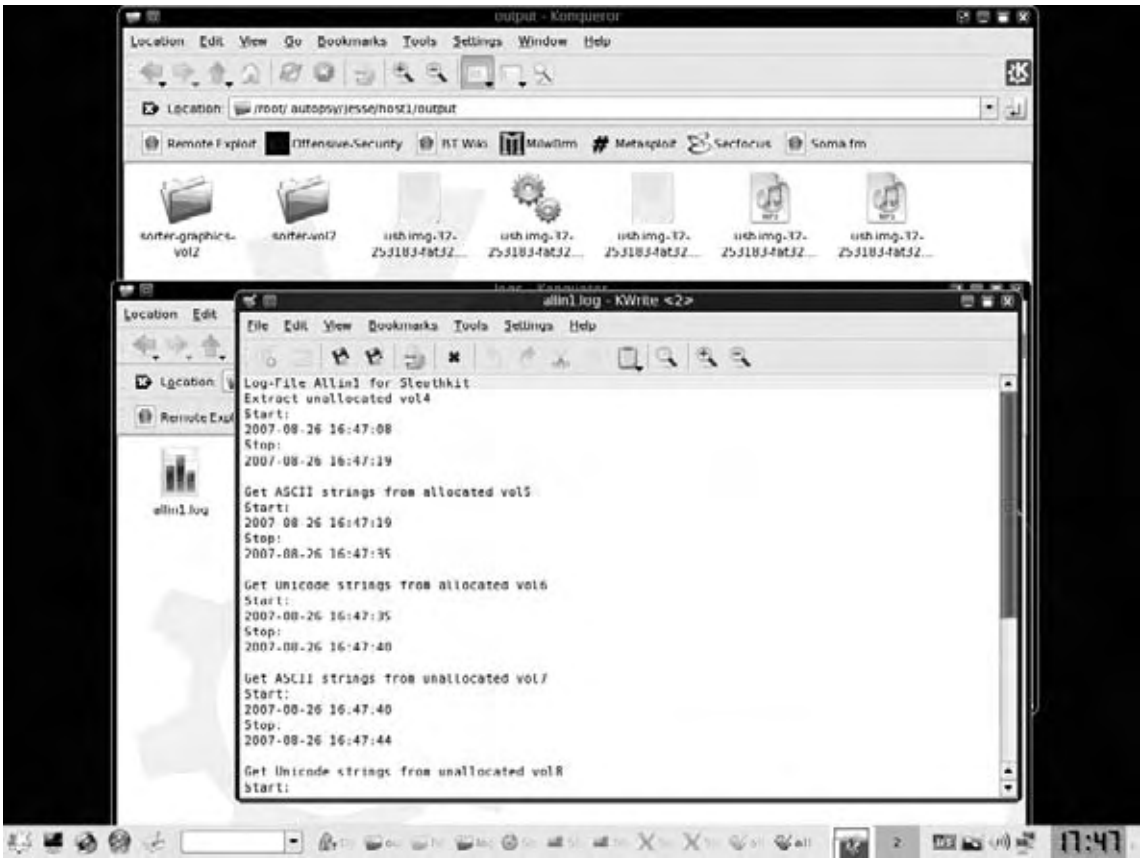
Using the Allin1 Tool

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI. Optional: It might be a good idea to make a disk image from a USB stick to do a practice run, preferably one that is smaller—say, 512-, 256-, or 128 MB with a FAT file system. Simply insert the USB stick, enter **fdisk –l** (should be SDA1 or similar), and then enter **dd if=/dev/sda of=/root/usb.img**.
2. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Autopsy**. (Starting Autopsy first will help to create the required host.aut file, which does not exist yet.)
3. Open the Konqueror or Firefox browser and go http://localhost:9999/autopsy. The screen will display Autopsy Forensic Browser 2.0 (refer back to Figure 11.12).
4. Click the **Open Case** radio button. If cases are present on the system, they will be listed.
5. Click the **New Case** radio button.
6. Give your case a name and description, add yourself as an investigator, and click the **New Case** radio button again.

Continued

7. Click the **Add Host** radio button and fill in the fields that you feel are pertinent to your investigation. Available fields include **Host Name**, **Description**, **Time Zone**, **Timeskew Adjustment**, **Path of Alert Hash Database**, and **Path of Ignore Hash Database**. Then click the **Add Host** radio button at the bottom of the screen. The configuration file host.aut has been created (see Figure 11.27). The location is /root/.autopsy/nameofcase/host1/host.aut, and it is needed for Allin1. (After you load the image file in steps 8–10, the host.aut file will be updated.)
8. Click **Add Image**.
9. Click **Add Image File**.
10. Under **Location:**, if you created a dd image of your USB stick earlier, it will be /root/usb.img. If not, select the path to your USB.
11. Select **Disk or Partition**. Under the Default import method, accept the default of **Symlink** and click **Next**. If the image was split, you would verify the split image files and click **Next** one additional time.
12. Close the Autopsy Browser. Open the Konqueror browser, and navigate to the /root directory. In the View menu, select **Show hidden files**. Double-click the **.autopsy** folder, double-click the **casename** folder, and double-click the **host1** folder. Notice that the output folder is empty, and in the log folder there is no Allin1 log.
13. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Allin1**.
14. For **Path to Sleuthkit binaries** (the default is incorrect), click **Choose** and select **/usr/local/sleuthkit/bin**.
15. For **Target**, click **Choose**. Right-click and select **Show hidden files**. Browse to /root/.autopsy/casename/host1/host.aut. Clicking the **Info** button should give you various details about the image file. If you do not receive details, something is wrong with the host.aut file.
16. Clicking **OK** will cause the screen to freeze until the analysis has been performed. The OK box will unfreeze and state **Finished** when the tool has completed all tasks.
17. Reopen the Konqueror browser and navigate to the /root directory. In the View menu, select **Show hidden files**. Double-click the **.autopsy** folder, double-click the **casename** folder, and double-click the **host1** folder. Look in the output folder for files, and look for and open the Allin1.log file in the log folder, as shown in Figure 11.28.

Figure 11.28 Output Folder, Contents of the Created Allin1.log File



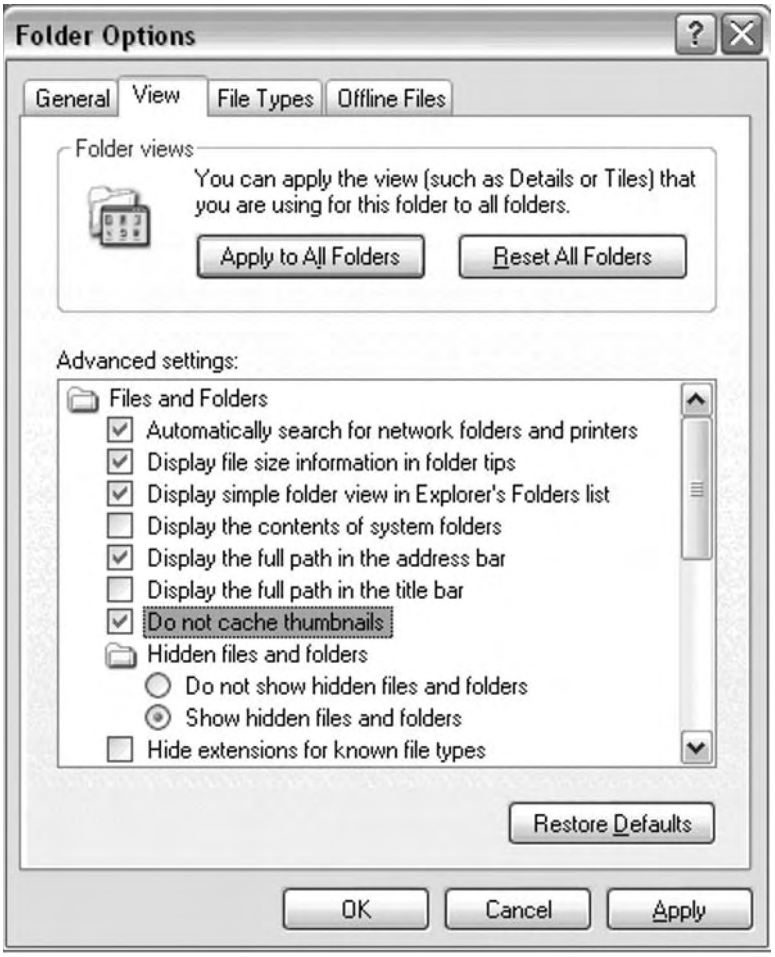
Vinetto

Vinetto was designed so that computer forensic investigators can extract information from thumbs.db files in Windows. If a Windows user selects the thumbnails view, a small database file called thumbs.db caches picture information to speed up the viewing of picture files within a folder.

NOTE

Windows Vista uses thumbcache.db instead of thumbs.db. Advanced users with systems running Windows XP, Windows 2000, or Windows 2003 may be aware of the forensic footprints left by thumbs.db files; these users can turn off thumbnail caching in Windows Explorer to prevent thumbs.db files from being created by using the Folder Options setting, as shown in Figure 11.29.

Figure 11.29 Preventing thumbs.db Files from Being Created


Folder Options
General
View
File Types
Offline Files
Folder views
You can apply the view (such as Details or Tiles) that you are using for this folder to all folders.
Apply to All Folders
Reset All Folders
Advanced settings:
Files and Folders
Automatically search for network folders and printers
Display file size information in folder tips
Display simple folder view in Explorer's Folders list
Display the contents of system folders
Display the full path in the address bar
Display the full path in the title bar
Do not cache thumbnails
Hidden files and folders
Do not show hidden files and folders
Show hidden files and folders
Hide extensions for known file types
Restore Defaults
OK
Cancel
Apply

If thumbs.db files are on a drive, you can use a simple *find* command in Linux to locate them and Vinetto to extract the metadata from them, as well as extract images and print an HTML report. The information you've extracted from a thumbs.db file can be very useful for determining what kinds of pictures are being stored on some Windows systems. The Vinetto tool is limited because certain versions of Windows are required, and thumbs.db files are not always present on examined systems.

Configuring & Implementing…

Using Vinetto to Extract Thumbnail Files

1. Boot to BackTrack on a system running a version of Microsoft Windows other than Vista. Log in as root with a password of toor, and type **startx** to bring up the GUI. Connect your destination media for extracted thumbs.db files USB stick or external drive
2. Select **K | BackTrack | Digital Forensics | Forensic Analysis | Vinetto**.
3. Type **find / -name Thumbs.db** (make sure you use a capital T).
4. If the files are located on the drive, you can use *vinetto –o* to extract the metadata (e.g., *vinetto –o "/mnt/hda1/Documents and Settings/Administrator/My Documents/My Pictures/Thumbs.db"*, as shown in Figure 11.30).
5. Type **vinetto –o** to extract the pictures within the thumbnail files. Make a directory on your removable drive by typing **mkdir /mnt/sda1/thumbs**; for example, *vinetto –o /mnt/sda1/thumbs "/mnt/hda1/Documents and Settings/Administrator/My Documents/My Pictures/Thumbs.db"*.
6. Type **vinetto –oH** to print out an HTML report to go along with the extracted pictures; for example, *vinetto –oH /mnt/sda1/thumbs "/mnt/hda1/Documents and Settings/Administrator/My Documents/My Pictures/Thumbs.db"*. The contents of your destination folder should resemble Figure 11.31.

Figure 11.30 Metadata Extracted from thumbs.db File Using Vinetto

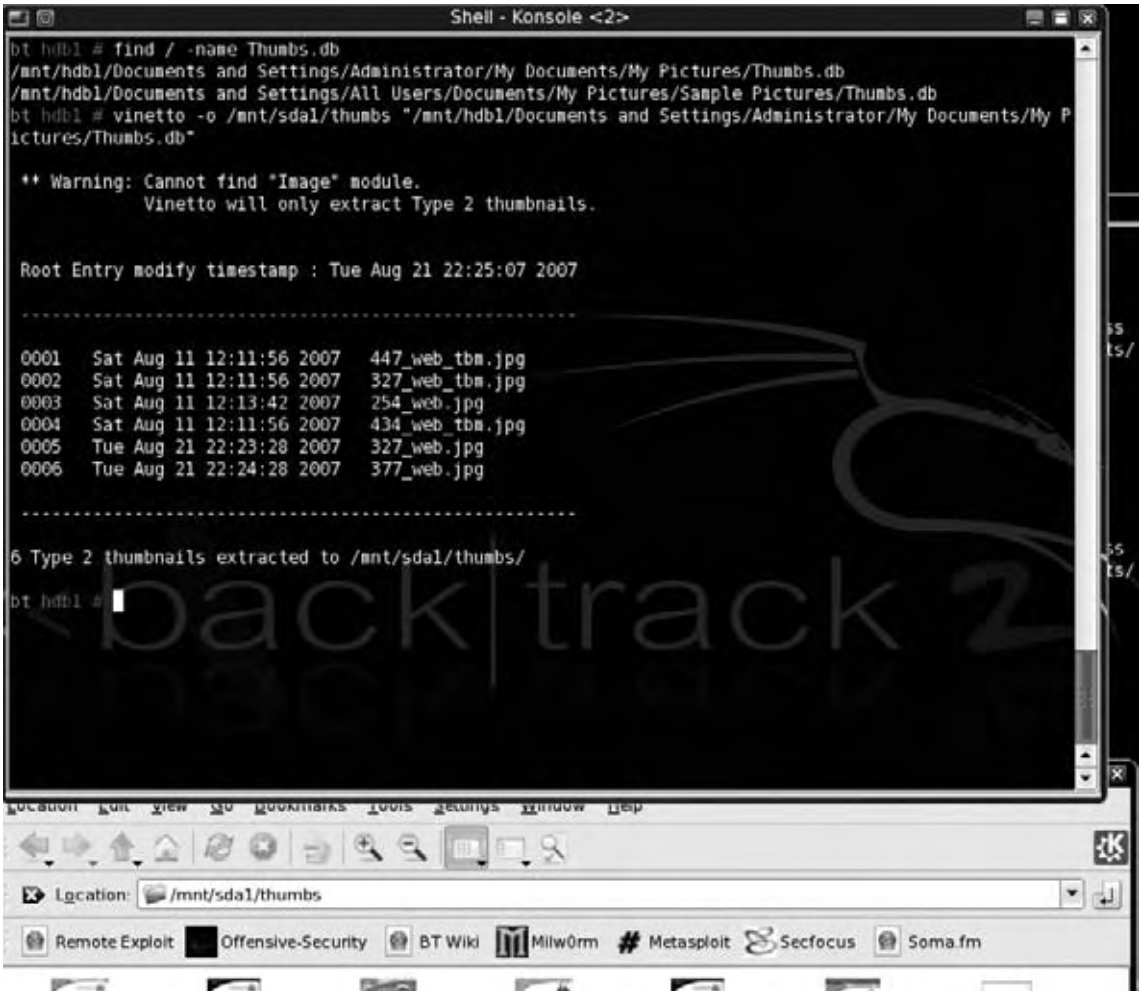
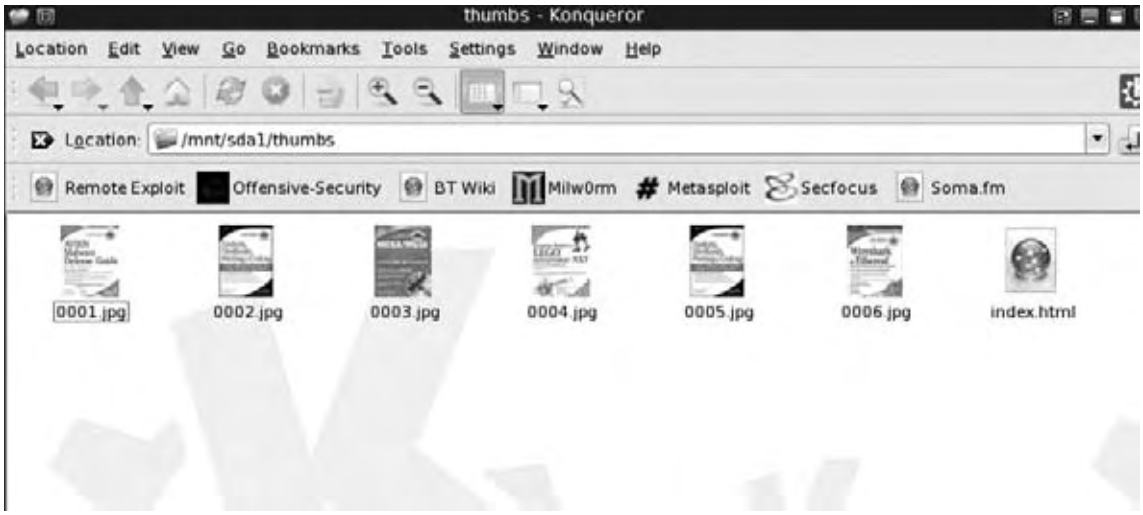

Shell - Konsole <2>
bt hdb1 # find / -name Thumbs.db
/mnt/hdb1/Documents and Settings/Administrator/My Documents/My Pictures/Thumbs.db
/mnt/hdb1/Documents and Settings/All Users/Documents/My Pictures/Sample Pictures/Thumbs.db
bt hdb1 # vinetto -o /mnt/sda1/thumbs "/mnt/hdb1/Documents and Settings/Administrator/My Documents/My P
ictures/Thumbs.db"
** Warning: Cannot find "Image" module.
Vinetto will only extract Type 2 thumbnails.
Root Entry modify timestamp : Tue Aug 21 22:25:07 2007
0001 Sat Aug 11 12:11:56 2007 447_web_tbm.jpg
0002 Sat Aug 11 12:11:56 2007 327_web_tbm.jpg
0003 Sat Aug 11 12:13:42 2007 254_web.jpg
0004 Sat Aug 11 12:11:56 2007 434_web_tbm.jpg
0005 Tue Aug 21 22:23:28 2007 327_web.jpg
0006 Tue Aug 21 22:24:28 2007 377_web.jpg
6 Type 2 thumbnails extracted to /mnt/sda1/thumbs/
bt hdb1 #
Location: /mnt/sda1/thumbs
Remote Exploit
Offensive-Security
BT Wiki
Milw0rm
Metasploit
Secfocus
Soma.fm

Figure 11.31 Thumbnail Pictures Extracted from thumbs.db file Using Vinetto



File Carving

File or data carving is the process of trying to reconstruct files based on their headers and footers, and on how the data is arranged. Files can be reconstructed from images, raw images, RAM, virtual memory, and file systems that have been damaged. Criminals sometimes change extensions on their files in an attempt to thwart detection by law enforcement or other people using the computer; indeed, file extensions can be changed on hundreds of files at once with the use of one simple command. Examiners use computer forensic software to determine whether suspects have used this technique in an attempt to cover their tracks. This process, known as file signature analysis, looks at the file headers to help identify the accurate file type. With file carving, when a certain file header is located in a stream of data, that header could help identify when the file begins.

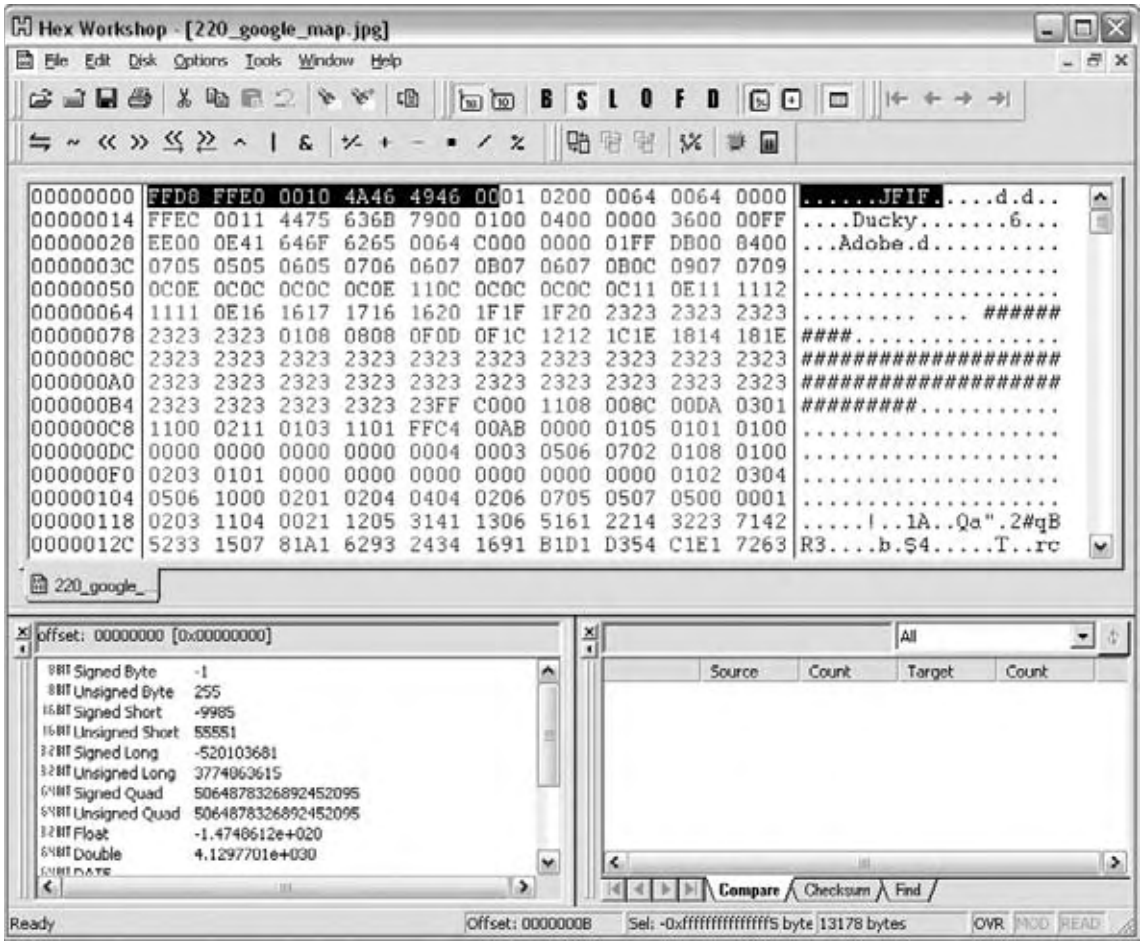
Notes from the Underground…

File Signatures

File signature headers are an accurate measure of file type. An extension is not necessarily a good indicator because a savvy suspect can change file extensions in an attempt to prevent detection. For example, the file funding.xls could be renamed to win32db.sys. Moving this file named win32db.sys into the system32 directory would make it even harder for someone to detect. Criminals often use the technique of renaming picture files to avoid detection from other users of the computer, or as an attempt to thwart law enforcement. You can use forensic tools to detect proper file types and help expose people who are trying to cover their tracks.

The file signature header and footer are extremely useful in assisting in the file-carving and recovery process. A hex view of a file shows the header and footer information, as shown in Figure 11.32. Gary Kessler, of Champlain College, has an extensive listing of common file headers on his Web site, www.garykessler.net/library/file_sigs.html. File carving involves finding the header and footer of a file in a stream of data, and using those two values to help extract and reconstruct the file. Although at times the file-carving process can be tedious and extremely time-consuming, some reconstructed files can provide investigators with a wealth of information regarding what kind of activity was conducted on a system. Foremost and Magicrescue are two valuable open source file-carving tools that are part of the BackTrack distribution. If you want to practice carving files or if you are up for a challenge, Digital Forensics Research Workshop, http://dfrws.org, provides raw data for file carving.

Figure 11.32 File Signature Header of a JPEG File



Foremost

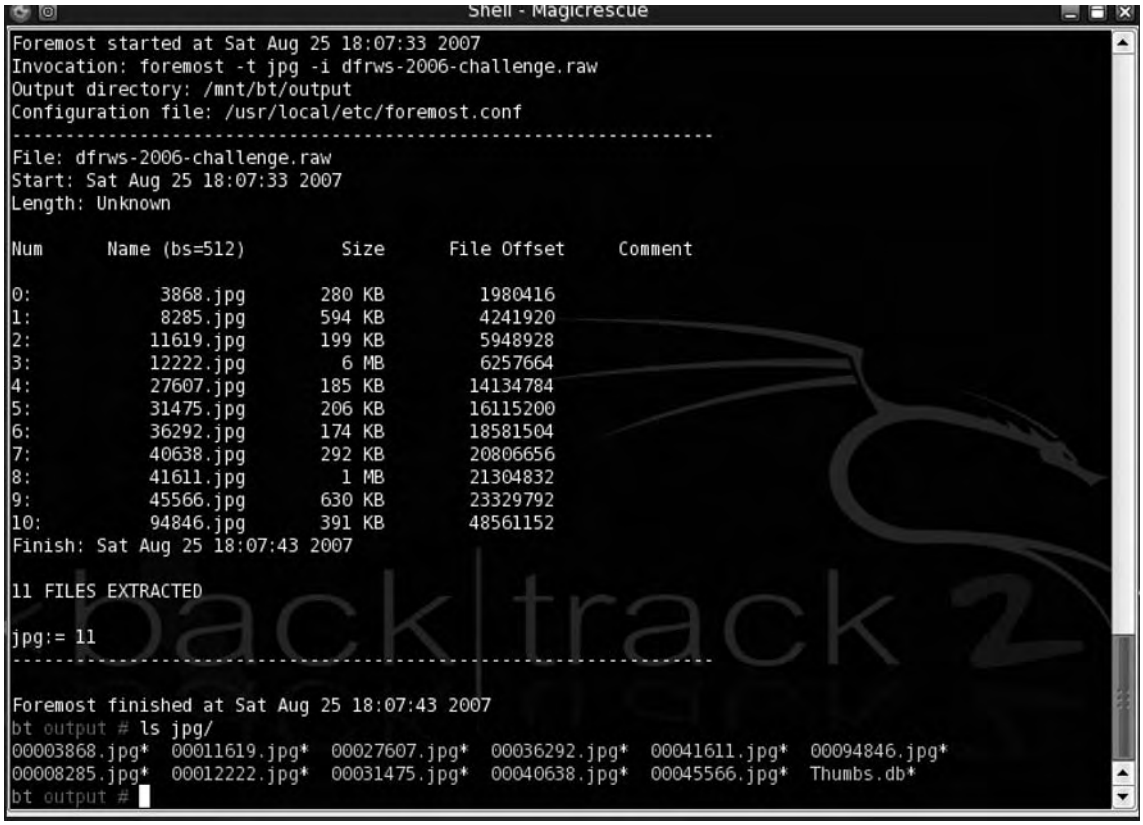
Foremost was developed by Jesse Kornblum, who works at the DoD Cyber Crime Center in Maryland. It is a command-line file-carving tool that will reconstruct files based on their headers, footers, and data format. When using Foremost, specify the type of file you are trying to recover and the location of the raw image file. After Foremost is finished, it creates a text file called audit.txt and an output folder where extracted files will be located. You do not have to have any knowledge of file carving to use the program; if you use the command with the appropriate switches, the existing files will be extracted automatically.

Configuring & Implementing…

Using Foremost to Carve Out JPEG Files

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
2. Open your Firefox browser and download http://dfrws.org/2006/challenge/dfrws-2006-challenge.zip to the /root directory.
3. Open a terminal and type **unzip dfrws-2006-challenge.zip**.
4. foremost –t jpg –i dfrws-2006-challenge.raw
5. After Foremost is finished, there will be a folder called output in the root directory (or whatever directory you ran the program from). Go into the folder, and notice that it contains an audit.txt file and another folder. The audit.txt file contains the information regarding what was extracted, and the folder—in this case, jpg—contains the reconstructed images of files based on the switch you used with Foremost (in this case, *foremost –t jpeg*). Read the text file by typing **more audit.txt**, as shown in Figure 11.33.
6. Go into the folder with all the recovered files, list the files, and open them with the appropriate viewer; for example, **cd jpg** or **kolorpaint 00011619.jpg**. Alternatively, open Konqueror and browse to the directory in which the files are listed and view them there.

Figure 11.33 Foremost audit.txt File and Recovered Files from an Image

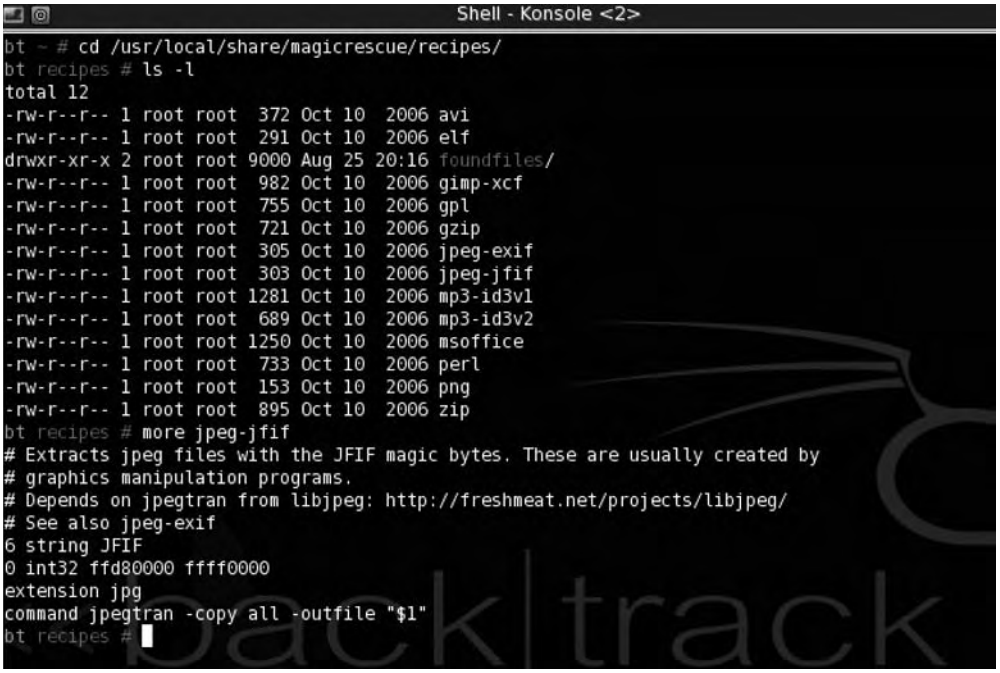

Shell - Magicrescue
Foremost started at Sat Aug 25 18:07:33 2007
Invocation: foremost -t jpg -i dfrws-2006-challenge.raw
Output directory: /mnt/bt/output
Configuration file: /usr/local/etc/foremost.conf
File: dfrws-2006-challenge.raw
Start: Sat Aug 25 18:07:33 2007
Length: Unknown
Num Name (bs=512) Size File Offset Comment
0: 3868.jpg 280 KB 1980416
1: 8285.jpg 594 KB 4241920
2: 11619.jpg 199 KB 5948928
3: 12222.jpg 6 MB 6257664
4: 27607.jpg 185 KB 14134784
5: 31475.jpg 206 KB 16115200
6: 36292.jpg 174 KB 18581504
7: 40638.jpg 292 KB 20806656
8: 41611.jpg 1 MB 21304832
9: 45566.jpg 630 KB 23329792
10: 94846.jpg 391 KB 48561152
Finish: Sat Aug 25 18:07:43 2007
11 FILES EXTRACTED
jpg:= 11
Foremost finished at Sat Aug 25 18:07:43 2007
bt output # ls jpg/
00003868.jpg* 00011619.jpg* 00027607.jpg* 00036292.jpg* 00041611.jpg* 00094846.jpg*
00008285.jpg* 00012222.jpg* 00031475.jpg* 00040638.jpg* 00045566.jpg* Thumbs.db*
bt output #

Magicrescue

Have you ever had your USB stick become corrupt or unreadable? If your file system is damaged, you can use Magicrescue to recover files from a block device. Magicrescue scans media for file types and sends the output to a folder. It uses magic bytes, or file signatures, to locate these files on the disk. Even if the user deleted files on the disk, if the data stream is still intact on the media, Magicrescue will extract the files.

BackTrack has 13 Magicrescue recipe files, shown in Figure 11.34, that have file signatures for JPEGs, MP3s, ZIPs, and several other file types. You can also create your own recipe file for a file type that is not listed among the 13 included with BackTrack. Running the Magicrescue program from the directory where the recipe files are located makes the syntax of the command a bit easier. I have used the program several times, and I have been amazed at the number of files the program recovered. When important files have been deleted or media have become damaged, try Magicrescue before you throw in the towel!

Figure 11.34 Magicrescue Recipe Files with "Magic Numbers"


Shell - Konsole <2>
bt ~ # cd /usr/local/share/magicrescue/recipes/
bt recipes # ls -l
total 12
-rw-r--r-- 1 root root 372 Oct 10 2006 avi
-rw-r--r-- 1 root root 291 Oct 10 2006 elf
drwxr-xr-x 2 root root 9000 Aug 25 20:16 foundfiles/
-rw-r--r-- 1 root root 982 Oct 10 2006 gimp-xcf
-rw-r--r-- 1 root root 755 Oct 10 2006 gpl
-rw-r--r-- 1 root root 721 Oct 10 2006 gzip
-rw-r--r-- 1 root root 305 Oct 10 2006 jpeg-exif
-rw-r--r-- 1 root root 303 Oct 10 2006 jpeg-jfif
-rw-r--r-- 1 root root 1281 Oct 10 2006 mp3-id3v1
-rw-r--r-- 1 root root 689 Oct 10 2006 mp3-id3v2
-rw-r--r-- 1 root root 1250 Oct 10 2006 msoffice
-rw-r--r-- 1 root root 733 Oct 10 2006 perl
-rw-r--r-- 1 root root 153 Oct 10 2006 png
-rw-r--r-- 1 root root 895 Oct 10 2006 zip
bt recipes # more jpeg-jfif
Extracts jpeg files with the JFIF magic bytes. These are usually created by
graphics manipulation programs.
Depends on jpegtran from libjpeg: http://freshmeat.net/projects/libjpeg/
See also jpeg-exif
6 string JFIF
0 int32 ffd80000 ffff0000
extension jpg
command jpegtran -copy all -outfile "$1"
bt recipes #

Configuring & Implementing…

Using Magicrescue to Find JPEG Files on a USB Thumb Drive

1. Boot to BackTrack. Log in as root with a password of toor, and type **startx** to bring up the GUI.
2. Open a terminal:

   ```
   cd /usr/local/share/magicrescue/recipes
   ls
   ```

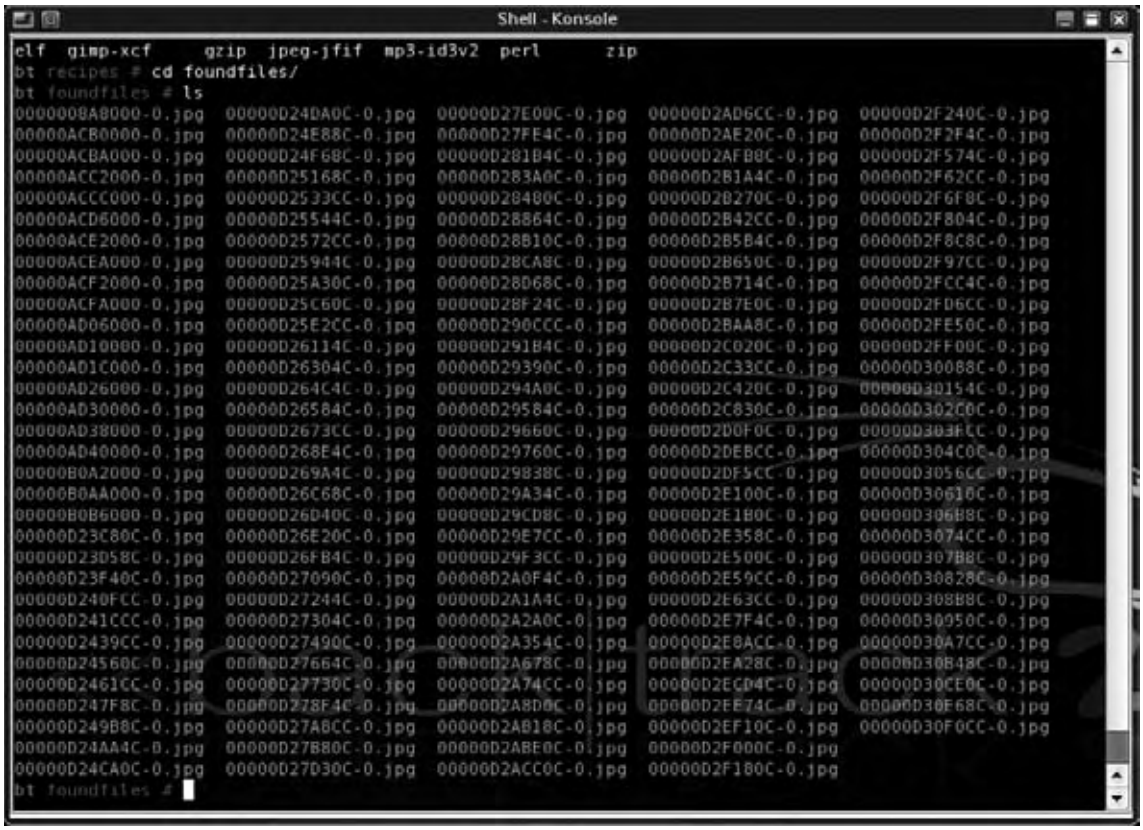
   This will list 13 predefined recipe files for the Magicrescue program.
3. Insert a USB thumb drive:

   ```
   fdisk –l (Should be /dev/sda1)
   mkdir /mnt/sda1/foundfiles
   ```

Continued

4. magicrescue –r jpeg-jfif –d /mnt/sda1/foundfiles /dev/sda1 (or whatever device name is)
5. cd /mnt/sda1/foundfiles && ls
6. kolorpaint name of jpg file.jpg (I found 158 jpeg files, as seen in Figure 11.35. (pretty impressive)

Figure 11.35 JPEGs Recovered by Magicrescue



Case Studies: Digital Forensics with the BackTrack Distribution



WARNING

Never make forensic copies of media without the owner's written consent. Advise your clients and customers to use login banners that warn system users that their activity is being monitored.

Digital forensics could be useful to a penetration tester in many situations. If systems have been hacked or a network intrusion has taken place, you can use forensic tools to examine volatile and nonvolatile data. You can use them to detect the presence of rootkits and malware. You can use the open source forensic tools included with BackTrack to make exact copies of media, including the media's slack space and deleted files and folders. You can extract pictures, e-mails, and documents from these images that can provide valuable clues regarding what types of activities were taking place on a system.

In the following scenario, you are contacted by a commercial client who is having bandwidth issues. After you arrive at the site, you conduct network monitoring and find an Internet Protocol (IP) address that is generating an unusual amount of traffic through port 80. You save some of your capture files and logs, write up a report, and inform the site director what is happening. The site director is not surprised by your findings; she has spoken with the individual who sits at that workstation on several occasions regarding his Internet usage. The site director asks you which sites the user has visited to try to gauge whether the user at that workstation is surfing Web sites related to his job. You tell the site director that all of the HTTP traffic from that user's station is directed to proxies and anonymizer sites. You explain to your client that individuals often use proxies and anonymizers to mask the actual sites they are visiting. The user from the workstation claims he is using that site for work-related issues only. The site director wants to find out what sites the user is visiting and what kinds of activity he is conducting during work hours.

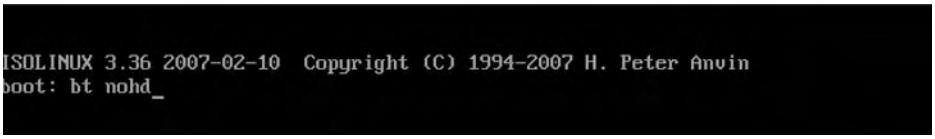
The site director has you meet with the network administrator. The admin tells you that she logged in to the workstation in question to check for cookies and browsing history and discovered that the user deletes his cookies and temporary Internet files at the end of every day. The network administrator also explains that the user seemed to know that someone logged in to the computer to investigate his behavior. When asked about the specifications of the company equipment, the admin says that all employees have the following:

- Dell OptiPlex GX240 2.0 GHz Pentium with 512 MB of RAM and a 40 GB hard drive running Windows XP Service Pack 2 with Internet Explorer Version 6 (the administrator installed the operating system to a single NTFS partition)
- Company-issued 512 MB Attache removable USB storage device

You explain to the director that you will be creating a forensic image of the user's hard drive and searching through the media for artifacts such as browser history and images. The systems are automatically shut off every night after backups have been made and updates have been installed. Here's what you should do:

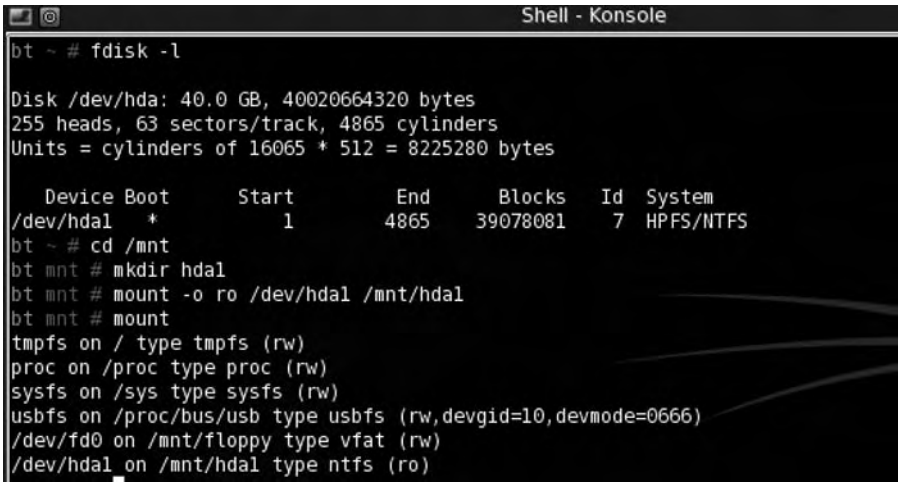
1. Boot to the user's workstation and enter the BIOS.
2. Document the current BIOS settings for the computer.
3. Set the CD-ROM to be the first boot device.
4. Exit the BIOS and save your changes.
5. Insert your BackTrack CD-ROM in the computer.
6. At the **boot:** prompt, type **bt nohd**, as seen in Figure 11.36. This will ensure that the drives are not auto-mounted.

Figure 11.36 Using the bt nohd Cheat Code So That Drives Will Not Be Auto-Mounted



7. Log in as root with a password of toor, and type **startx** to bring up the GUI. Then open a terminal session and type **fdisk –l**.
8. You will see the drive letter and numbering designation of the source disk, as seen in Figure 11.37. (This is most likely HDA1 for a single IDE drive with one partition.)

Figure 11.37 Mounting the Source Disk As Read-Only


Shell - Konsole
bt ~ # fdisk -l
Disk /dev/hda: 40.0 GB, 40020664320 bytes
255 heads, 63 sectors/track, 4865 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 4865 39078081 7 HPFS/NTFS
bt ~ # cd /mnt
bt mnt # mkdir hda1
bt mnt # mount -o ro /dev/hda1 /mnt/hda1
bt mnt # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/fd0 on /mnt/floppy type vfat (rw)
/dev/hda1 on /mnt/hda1 type ntfs (ro)

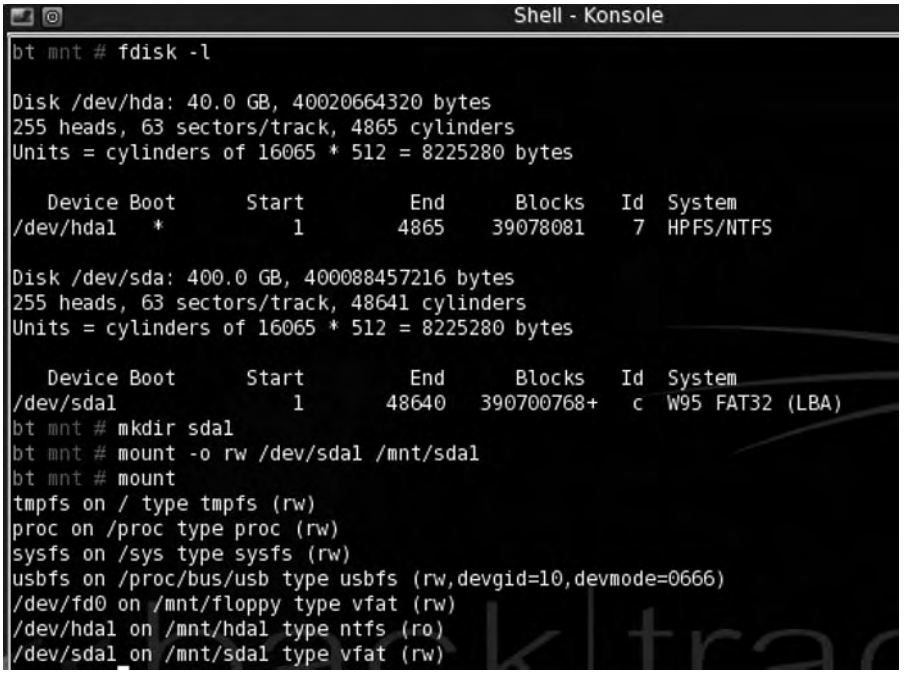
9. Mount the drive as read-only, as seen in Figure 11.37:

```
ex:
cd /mnt
mkdir hda1
mount –o ro /dev/hda1 /mnt/hda1
```

10. Plug in the USB mass storage device.
11. Type **fdisk –l** again. You should see a new, added drive letter and partition, most likely SDA1 (as seen in Figure 11.38). This is your destination disk.
12. Mount this drive as read/write, as shown in Figure 11.38:

```
ex:
cd /mnt
mkdir sda1
mount –o rw /dev/sda1 /mnt/sda1
```

Figure 11.38 Mounting the Destination USB Mass Storage Drive As Read/Write


Shell - Konsole
bt mnt # fdisk -l
Disk /dev/hda: 40.0 GB, 40020664320 bytes
255 heads, 63 sectors/track, 4865 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 4865 39078081 7 HPFS/NTFS
Disk /dev/sda: 400.0 GB, 400088457216 bytes
255 heads, 63 sectors/track, 48641 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sda1 1 48640 390700768+ c W95 FAT32 (LBA)
bt mnt # mkdir sda1
bt mnt # mount -o rw /dev/sda1 /mnt/sda1
bt mnt # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/fd0 on /mnt/floppy type vfat (rw)
/dev/hda1 on /mnt/hda1 type ntfs (ro)
/dev/sda1 on /mnt/sda1 type vfat (rw)

13. Make a directory on the destination disk, called casestudy, and go into that directory by typing **mkdir casestudy && cd /mnt/sda1/casestudy**.

Now you should use Vinetto to try to recover images from thumbs.db files that the user may have on his system, and pipe the images to the destination USB drive:

```
find / -name Thumbs.db
```

Unfortunately, there are no hits, as seen in Figure 11.39. It is possible that this user was aware of the cache left by thumbnails and disabled it. This is a good example of why you may need to use more than one forensic tool, depending on the circumstances and configurations.

Figure 11.39 No Results from the Search for thumbs.db Files



If a thumbnail cache had existed, you could have easily redirected it to the external USB drive:

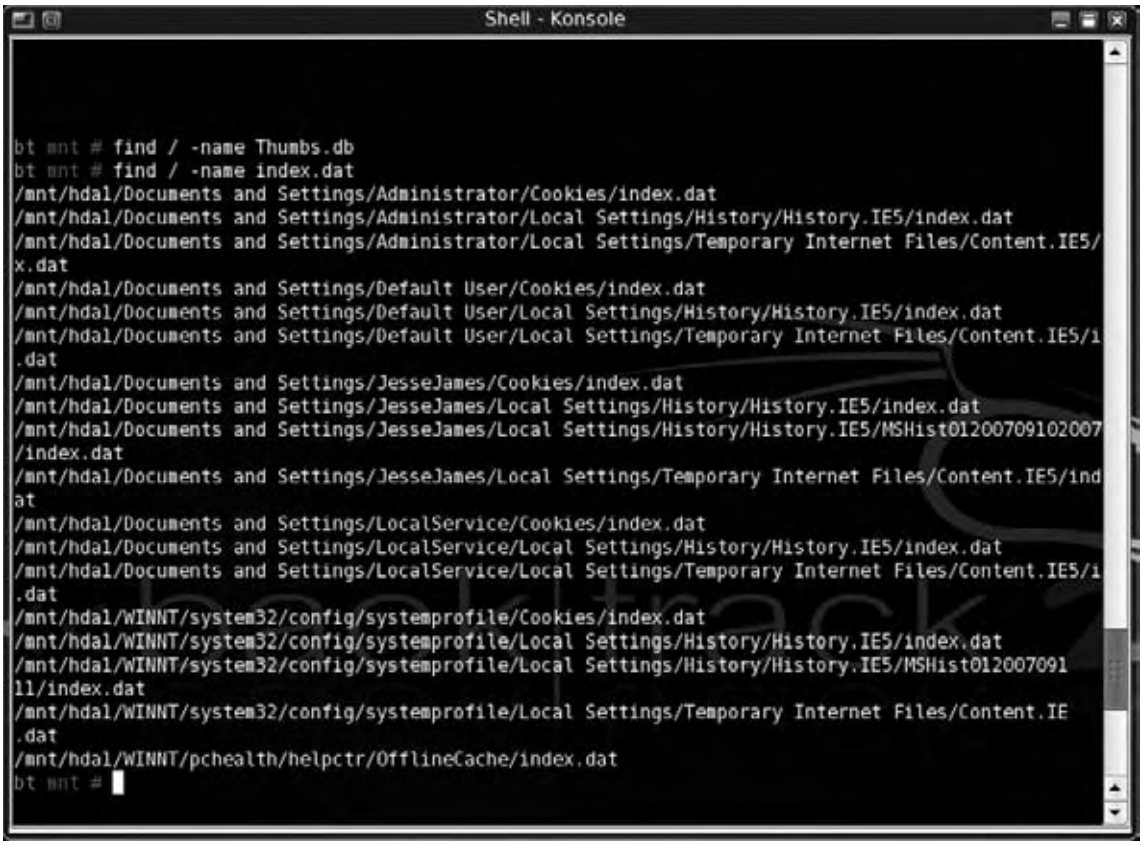
```
ex:
mkdir /mnt/sda1/casestudy/thumbs
vinetto -oH /mnt/sda1/casestudy/thumbs "/mnt/hda1/Documents and Settings/
JesseVarsalone/My Documents/My Pictures/Thumbs.db"
```

Now you should try to look for the index.dat files, which may wind up giving you some more detailed information regarding some of the Web sites the user has browsed. You can use the pasco tool to read the index.dat files:

```
find / -name index.dat
```

This time you get several search results, as seen in Figure 11.40.

Figure 11.40 Results from the Search for index.dat Files



You need to use the following syntax for pasco (include the full path to the file in quotes):

```
pasco "/mnt/hda1/Documents and Settings/JesseJames/UserData/index.dat"
```

To redirect output into a text file, as seen in Figure 11.41, use the *–d* switch with pasco to uncover even more information about the user's browsing (note that copying and pasting works very well here):

```
pasco –d "/mnt/hda1/Documents and Settings/JesseJames/UserData/index.dat" > /mnt/
sda1/casestudy/browserhistory1.txt
```

In this case, all four files related to the browser activity of this user, Jesse James, were redirected to the USB mass storage device. Even though results were produced in this case, as seen in Figure 11.40, if the user was using a browser other than Internet Explorer, this tool would probably not have been helpful. In Figure 11.41, you can view the results from one of the four index.dat files. Within the information recovered in the index.dat files, there were several indications that the user was going to poker sites and trying to book a room in Las Vegas for the upcoming weekend. Figure 11.41 shows part of the browsing information

pasco provided when reading the index.dat file. Notice that the URLs the user visited are listed along with the time and date stamps.

Figure 11.41 A Partial Screen Shot of One of Four index.dat Files

```
History File: /mnt/hda1/Documents and Settings/JesseJames/Local
Settings/History/History.IE5/index.dat

TYPE URL MODIFIED TIME ACCESS TIMEFILENAME DIRECTORY
HTTP HEADERS
URL Visited: JesseJames@http://anonimzer.com Tue Sep 11 05:32:56 2007
Tue Sep 11 05:32:56 2007 URL
URL Visited: JesseJames@about:Home Tue Sep 11 04:59:17 2007 Tue
Sep 11 04:59:17 2007 URL
URL Visited:
JesseJames@http://www.boston.com/news/globe/living/articles/2007/09/10/timberlake_rihanna_r
ack_em_up_in_las_vegas Tue Sep 11 05:56:56 2007 Tue Sep 11 05:56:56 2007
URL
URL Visited:
JesseJames@http://adserving.cpxinteractive.com/rw?title=New%20offer%21&qs=iframe3%3FA
AAAALnRAQBLIQUAvOQBAAIAAAAAAP8AAAACFAICAAL60gEAuwsDAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAADgic%2DfF0AAAAAAAAA
AAAAAwH0pLCJAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAghQ73pa4LgP129WC4wl%2El3ghaQa2NRIAP4BPsQA
AAAA%3D%2C%2Chttp%3A%2F%2Fanonimzer%2Ecom%2Fpp%2Ephp%3Fpid%3D Tue
Sep 11 05:32:58 2007 Tue Sep 11 05:32:58 2007 URL
URL Visited: JesseJames@http://www.google.com Tue Sep 11 06:00:56 2007
Tue Sep 11 06:00:56 2007 URL
URL Visited: JesseJames@http://www.freeproxy.ru/en/free_proxy/cgi-proxy.htm
Tue Sep 11 05:33:22 2007 Tue Sep 11 05:33:22 2007 URL
URL Visited: JesseJames@http://www.anonymizer.com Tue Sep 11 05:33:12
2007 Tue Sep 11 05:33:12 2007 URL
URL Visited: JesseJames@http://www.google.com/search?hl=en&q=anonymizer
Tue Sep 11 05:33:38 2007 Tue Sep 11 05:33:38 2007 URL
URL Visited: JesseJames@http://cexx.org/anony.htm Tue Sep 11 05:35:34 2007
Tue Sep 11 05:35:34 2007 URL
URL Visited: JesseJames@http://www.google.com/search?hl=en&q=free+proxies
Tue Sep 11 05:34:00 2007 Tue Sep 11 05:34:00 2007 URL
URL Visited: JesseJames@http://www.proxy4free.com/page1.html Tue Sep 11
05:34:03 2007 Tue Sep 11 05:34:03 2007 URL
URL Visited: JesseJames@http://www.vegas.com Tue Sep 11 05:54:31 2007
Tue Sep 11 05:54:31 2007 URL
URL Visited: JesseJames@http://www.google.com/search?hl=en&q=poker Tue
Sep 11 05:55:10 2007 Tue Sep 11 05:55:10 2007 URL
URL Visited: JesseJames@http://www.pokerroom.com/poker/tournaments.html
Tue Sep 11 05:57:20 2007 Tue Sep 11 05:57:20 2007 URL
URL Visited: JesseJames@http://www.pokerroom.com Tue Sep 11 05:55:57
2007 Tue Sep 11 05:55:57 2007 URL
```

Now, you need to make a forensic copy of the disk. Hash the source drive, as seen in Figure 11.42:

```
md5sum /dev/hda
```

Figure 11.42 Hashing the Source Disk



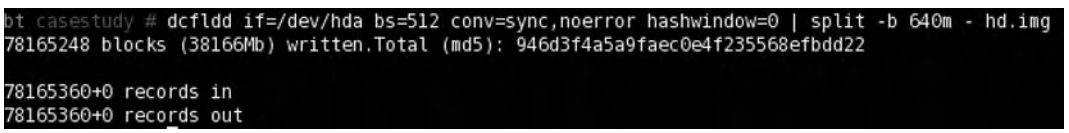
The results of the hash can also be piped to the output to the USB destination disk:

```
md5sum /dev/hda > /mnt/sda1/casestudy/hash.txt
```

Use dcfldd to create a split image on your USB mass storage device, as seen in Figure 11.43. The *hashwindow=0* switch will return the MD5 hash of the disk:

```
dcfldd if=/dev/hda bs=512 conv=sync,noerror hashwindow=0 | split –b 640m – hd.img
```

Figure 11.43 Using the DCFLDD Command to Create a Forensic Image

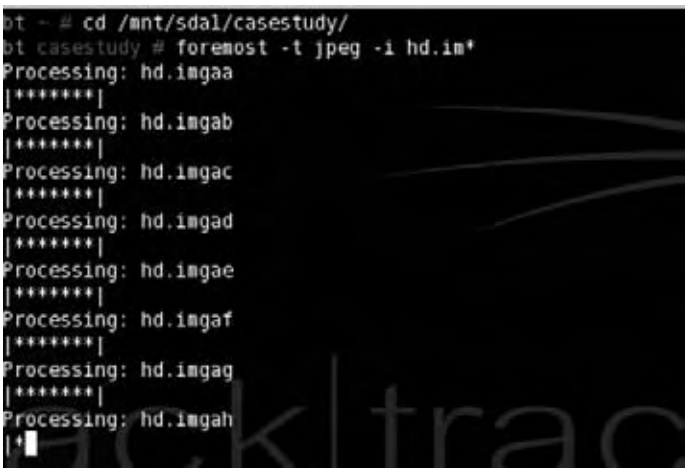

bt casestudy # dcfldd if=/dev/hda bs=512 conv=sync,noerror hashwindow=0 | split -b 640m - hd.img
78165248 blocks (38166Mb) written.Total (md5): 946d3f4a5a9faec0e4f235568efbdd22
78165360+0 records in
78165360+0 records out

Notice that the two hash values match in Figure 11.42 and Figure 11.43. You can be fairly confident that the data sets are the same; the chance of two file s having the same MD5 hash is 1 in 2^{128}. Disconnect the hard drive and restart the system by opening the terminal and typing: **shutdown –r now**. Restore the BIOS changes you made earlier to the settings you properly documented. Turn off the workstation, and connect the USB drive to your forensic workstation (or another computer you own). Boot the computer up to BackTrack.

Now use the Foremost tool to carve some images from the raw dd image you created using dcfldd. The image was split, so you should use * after hd.img to include all the image files, as seen in Figure 11.44:

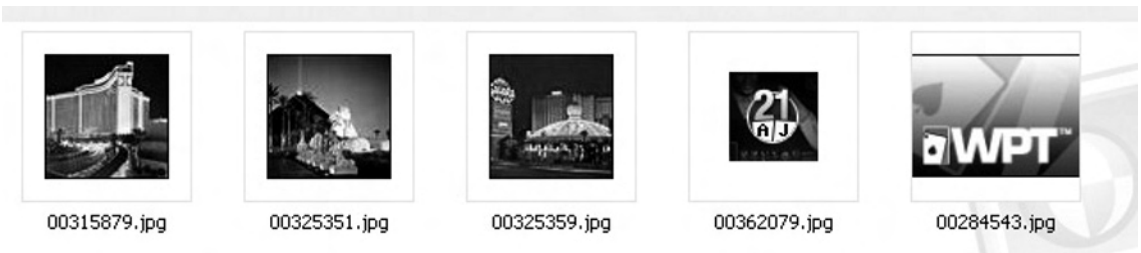
```
cd /mnt/sda1/casestudy
foremost –t jpeg –i hd.img*
```

Figure 11.44 Using the Foremost Tool to Data-Carve File Images



The Foremost tool worked very well, and it recovered more than 1,000 images. Many of the images were part of the Windows operating system; do not be surprised if Foremost recovers images that are several years old. Use Konqueror to browse through the images. Using Ksnapshot, take some relevant images to share with the site director. 11If you see anything that looks illegal, contact the proper authorities. Figure 11.45 shows a sample of some of the poker site pictures that were found.

Figure 11.45 Pictures Extracted from the Image of the Hard Drive Using Foremost



At this point, you present your evidence to your client. When confronted with the evidence, the employee is terminated and must return the company's USB drive. He tells the site director not to bother looking for files on it because he used a tool to format it right before he was called into the office.

This is where you will use the Magicrescue tool to see whether any images from the USB drive are recoverable:

1. Insert your BackTrack CD-ROM into the computer.
2. At the boot: prompt, type **bt nohd**.
3. Insert the USB stick.
4. Open a terminal:

```
fdisk -l (Should be /dev/sda1)
```

5. Mount the device as read-only, as seen in Figure 11.46:

```
cd /mnt
mkdir sda1_removable
mount -o ro /dev/sda1 /mnt/sda1_removable
```

6. Plug in the USB mass storage device.

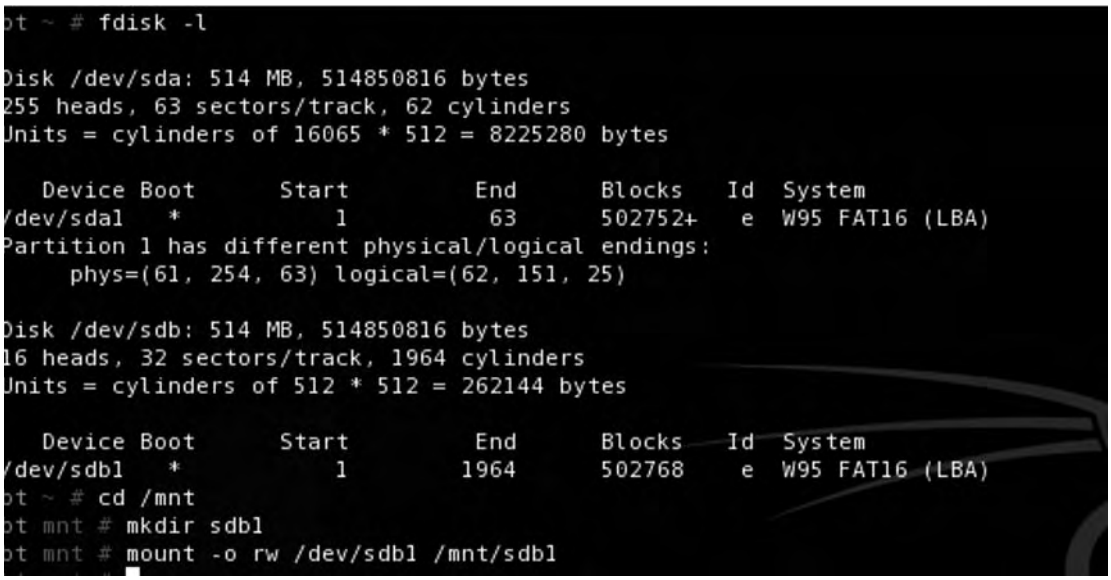
Figure 11.46 Mounting the Destination USB Mass Storage Drive As Read/Write


bt ~ # fdisk -l
Disk /dev/sda: 514 MB, 514850816 bytes
255 heads, 63 sectors/track, 62 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 63 502752+ e W95 FAT16 (LBA)
Partition 1 has different physical/logical endings:
phys=(61, 254, 63) logical=(62, 151, 25)
bt ~ # cd /mnt
bt mnt # mkdir sda1_removable
bt mnt # mount -o ro /dev/sda1 /mnt/sda1_removable/
bt mnt # mount
tmpfs on / type tmpfs (rw)
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
usbfs on /proc/bus/usb type usbfs (rw,devgid=10,devmode=0666)
/dev/sda1 on /mnt/sda1_removable type vfat (ro)

7. Type **fdisk –l** again. You should see a new, added drive letter and partition, most likely sdb. (Note that the thumb drive is the first device, so the USB mass storage device will be the second.) This is your destination disk. This drive needs to be mounted as read/write, as shown in Figure 11.47:

```
ex:
cd /mnt
mkdir sdb1
mount -o rw /dev/sdb1 /mnt/sdb1
```

Figure 11.47 Mounting the Destination USB Mass Storage Drive As Read-Write


bt ~ # fdisk -l
Disk /dev/sda: 514 MB, 514850816 bytes
255 heads, 63 sectors/track, 62 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sda1 * 1 63 502752+ e W95 FAT16 (LBA)
Partition 1 has different physical/logical endings:
phys=(61, 254, 63) logical=(62, 151, 25)
Disk /dev/sdb: 514 MB, 514850816 bytes
16 heads, 32 sectors/track, 1964 cylinders
Units = cylinders of 512 * 512 = 262144 bytes
Device Boot Start End Blocks Id System
/dev/sdb1 * 1 1964 502768 e W95 FAT16 (LBA)
bt ~ # cd /mnt
bt mnt # mkdir sdb1
bt mnt # mount -o rw /dev/sdb1 /mnt/sdb1

8. Browse to the casestudy directory on the USB mass storage device:

```
cd /mnt/sdb1/casestudy
mkdir foundfiles
```

9. Go into the Magicrescue recipes directory.
10. Pipe the output of Magicrescue to your removable USB mass storage device, as seen in Figure 11.48:

```
cd /usr/local/share/magicrescue/recipes
magicrescue –r jpeg-jfif –d /mnt/sdb1/casestudy/foundfiles /dev/sda1
(or whatever device name is)
cd foundfiles && ls
```

Figure 11.48 Piping the Output of Magicrescue to the USB Mass Storage Device

```
bt sdb1 # mkdir casestudy
bt sdb1 # cd casestudy
bt casestudy # mkdir foundfiles
bt casestudy # cd /usr/local/share/magicrescue/recipes/
bt recipes # magicrescue -r jpeg-jfif -d /mnt/sdb1/casestudy/foundfiles/ /dev/sda1
Found jpeg-jfif at 0x42000
/mnt/sdb1/casestudy/foundfiles//000000042000-0.jpg: 3661 bytes
Found jpeg-jfif at 0x64A000
/mnt/sdb1/casestudy/foundfiles//00000064A000-0.jpg: 3908 bytes
Found jpeg-jfif at 0x64C000
/mnt/sdb1/casestudy/foundfiles//00000064C000-0.jpg: 4847 bytes
Found jpeg-jfif at 0x64E000
/mnt/sdb1/casestudy/foundfiles//00000064E000-0.jpg: 4022 bytes
Found jpeg-jfif at 0x650000
/mnt/sdb1/casestudy/foundfiles//000000650000-0.jpg: 2168 bytes
Found jpeg-jfif at 0x652000
/mnt/sdb1/casestudy/foundfiles//000000652000-0.jpg: 4767 bytes
Found jpeg-jfif at 0x8A8000
/mnt/sdb1/casestudy/foundfiles//0000008A8000-0.jpg: 30516 bytes
```

11. Use Konqueror to browse through the images. Using Ksnapshot, take some relevant images to share with the site director. Figure 11.49 shows a sample of some of the poker site pictures you were able to recover.

Figure 11.49 Magicrescue-Recovered JPEG Images from USB Thumb Drive



Summary

The open source forensic tools included with the BackTrack distribution are quite powerful. They can help you in a variety of tasks, such as recovering deleted files, analyzing thumbnails and index.dat files, and capturing volatile and nonvolatile storage. Digital forensics is an important component of computer security; the scope of computer forensics continues to expand and its use is becoming more pertinent to companies around the globe. You can use the tools in BackTrack in a variety of ways to investigate incidents that take place within an IT infrastructure.

Chapter 12

Windows, Linux, and Macintosh Boot Processes

Solutions in this chapter:

- **The Boot Process**
- **The Macintosh Boot Process**
- **Macintosh Forensic Software**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

In this chapter we will look at the startup process that most computers go through and how the fundamental operating systems get loaded and started. You will see that computers start with tiny steps that build on each other, getting larger until the entire system is loaded and running. Only then can you, the end user, issue commands that the computer interprets and understands.

One of the most popular analogies for how a computer starts up is the amnesia scenario. For a moment look around you at the things you use everyday: telephones, pencils, coffee cups, and so on. Now imagine that you closed your eyes and when you opened them you didn't recognize any of those things, and didn't know how they worked. That is what happens inside a computer when you press the reset or the power button.

At the most fundamental level, computers understand only two things: true and false. The process of getting the computer from being a completely blank state to a fully running operating system is one of the fundamental items that every investigator should understand.

After looking at how a Macintosh boots, we will look at some of the tools that are available for analyzing Macintosh systems using both the Macintosh and Windows operating systems.

The Boot Process

The term "boot," depending on whom you talk to, came either from the old phrase, "Pulling one's self up by the bootstraps," or just from the word "bootstrap," meaning the leather tabs you use to pull on your boots. Either way it is a part of computer history and lore and is commonly used as the computer term for the initial startup of the system. All systems that are able to run Microsoft or Linux operating systems use the same boot up process. Once the computer completes this initial startup the specific operating system will load what it needs to continue. First we will look at the boot process in detail.

System Startup

When the computer power is first applied, a series of very small programs are executed in series leading to the eventual fully running system. Each small program builds on the previous until the entire operating system is loaded and operational.

Much like the character in the movie where the person with amnesia wakes up every morning not knowing what happened the day before, the computer needs to rediscover the world every time it is awakened.

When the power switch is engaged, circuits in the power supply go through an initial surge and then settle to a smooth output of 5 and 12 volts typically. When the voltage and amperage are within specifications, a power-good signal is generated and sent to the microprocessor. Think of this as the computer's wake-up call.

POST: Power On Self Test

The microprocessor then goes through a reset that clears its registers and memory and begins to execute the program stored in the Read-only Memory (ROM) Basic Input-Output System (BIOS) starting at address hF000. This small program tests the basic motherboard functions and will report any errors with audio beep codes. This small program is known as the Power On Self Test (POST).

In addition to the audio beeps, most modern computers also send a diagnostic code to internal port h80 if an error occurs. Diagnostic cards are available that plug into the ISA/PCI/AGP slots on motherboard that read these codes and give more comprehensive error reports.

The next step is for the BIOS to search for a video display adapter that has its own ROM. After verifying the checksum of the video ROM the BIOS will execute that ROM code to initialize the video adapter. Successful completion of the video ROM test and load usually produces a blinking cursor on the screen at a minimum. Different video ROMs do different things on completion, including flashy rainbow colored logos, for example.

The BIOS then checks the value at memory location h0472 for the existence of h1234, which is the indicator of a warm start, meaning that the computer is being restarted. Any other value indicates the computer is coming up from a cold start or powered off state.

If the computer is being cold started then the complete BIOS POST is performed, otherwise an abbreviated POST is done. In most modern computers the BIOS settings will let you skip things like full memory testing if you are just restarting the machine.

If a cold start is being performed the computer will read and write to every memory location in Random Access Memory (RAM) to verify that it is working properly. On most computer systems, you will see a counter or progress bar on the screen at this stage showing the status of this memory test.

It is at this point that the POST looks for the existence of a keyboard, and checks for any key press actions. If keys have been pressed, the program stores that information for action at the completion of the POST process. The keys pressed may be requests for access to the BIOS setup menu or the key command to boot from the network.

The computer then checks for the existence of any floppy or hard disk drives. Any devices found are compared to the contents of the configuration information stored in the CMOS memory chip on the motherboard. This chip has a small battery attached to it, keeping the data stored inside active when the computer is powered off.

The BIOS then scans memory for any other ROMs in the range of hC800 to hDF80 that may contain any additional code needing to be executed for other adapter cards like sound cards, RAID arrays, and such. If a ROM is found then the checksum is tested and programs found are run for each device in order.

If new devices are found during this stage the system may display the BIOS Setup screen, allowing you to make the modifications permanent.

Traditionally, if you hear a single beep, the system has completed the POST and is proceeding with Operating System Load.

The Master Boot Record

The BIOS will then search for a volume boot sector at cylinder 0, head 0, and sector 1 of the default boot drive. Historically the default boot drive was the first floppy disk drive. Modern configurations allow you to designate any drive, or external device like a USB thumb drive or even the network interface card as a boot device.

If there doesn't appear to be bootable code at the first sector, then the BIOS will look for a Master Partition Record at that location. If there is no Master Partition Record then the system will display an error similar to:

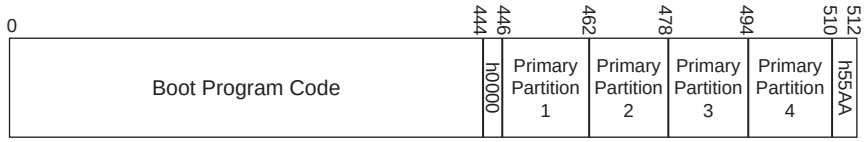
> Non-System disk or disk error
>
> Replace and strike any key when ready
>
> DISK BOOT FAILURE, INSERT SYSTEM DISK AND PRESS ENTER

If a volume boot record is found, the program is loaded and tested to ensure that the last two bytes are h55AA, and then the program is run. This program is very small since it must stay within a 512 byte sector. The actual program can only be 444 bytes long.

The Master Boot Record (MBR) is traditionally the first 512 bytes, or the first sector (sector 0), of a storage device like a hard disk drive that is bootable.

The 512 bytes are broken down as shown in Figure 12.1; the first 444 bytes are allocated for any boot program, end of code marker of two nulls (h00), four 16 byte master partition records, and the end of MBR marker of h55AA.

Figure 12.1 The MBR Layout



Master Boot Record

The Master Partition Table entries are broken down as shown in Table 12.1. There can be up to four entries in this area of the MBR.

Table 12.1 Master Partition Table Entries

Byte Offset/Length	Contents	Value
0 (I byte)	Boot Indicator Byte	h80 = Bootable h00 = Nonbootable
1 (3 bytes)	Cylinder/Head/Sector (CHS) of Partition Start	(Variable)
4 (1 byte)	Partition Type	h00 = No Partition h01 = DOS 12 bit FAT h04 = DOS 16 bit FAT h05 = Extended DOS h06 = Primary DOS h07 = OS2 h08 = AIX (File Sys) h09 = AIX (boot) h64 = Novell NetWare hFF = Unix Bad Block Table
5 (3 bytes)	Cylinder/Head/Sector (CHS) of Partition End	(Variable)
8 (4 bytes)	Relative Sector Offset of Partition	(Variable)
12 (4 bytes)	Partition size in sectors	(Variable)

The Partition Type values shown are some of the more common found; different operating systems can use different values.

Damage & Defense…

The Worst Place to Get a Virus

In the days of bootable floppy disks, Master Boot Record (MBR) viruses were much more prevalent. With fewer people using or even having a floppy disk drive in their computer, MBR attacks are harder to accomplish. This does not mean that they can't be done!

It is a very good practice to scan boot media that is from a less than obvious source. Placing MBR viruses on bootable USB media or copied CDs and DVDs is still a very real threat.

Since MBR code runs well before any antivirus program it can infect and rootkit your machine before you even realize it.

Loading MSDOS

The volume boot record program has only two things it must do: look for the existence of the files, IO.sys and MSDOS.sys (or in some cases IBMDOS.com); and then load and run IO.sys. Once IO.sys is running, the boot record program is no longer needed.

IO.sys contains a routine called SYSINT, which facilitates the loading of additional drivers and settings that the system will need in order to complete loading the operating system.

Part of the SYSINT process loads the MSDOS.sys program that resets the disk system, resets and initializes attached devices, and sets the system to the default parameters.

If SYSINT finds a CONFIG.sys file in the root directory of the default boot device, it passes the contents to MSDOS.sys for processing. The CONFIG.sys file contains directives that define the number of files and file buffers to allocate for this system, and may contain more device drivers to be loaded for special attached devices.

Then the SYSINT routine loads COMMAND.com, which is referred to as a shell program. It contains the last of the device extensions to the base operating system. It also contains the standard command line programs like ECHO, DIR, and FORMAT.

The last thing that COMMAND.com has SYSINT do is look for a file in the root directory of the default drive called AUTOEXEC.bat, which may contain programs and commands that the user wants executed each time at startup. If this file is found, COMMAND.com executes the directives in the file and when that is complete the user is presented with the command prompt. The default command prompt is the default drive letter, a colon, and the greater than sign (example: "A:>"). This prompt can be changed by the Prompt command in the AUTOEXEC.bat file.

Loading Windows XP

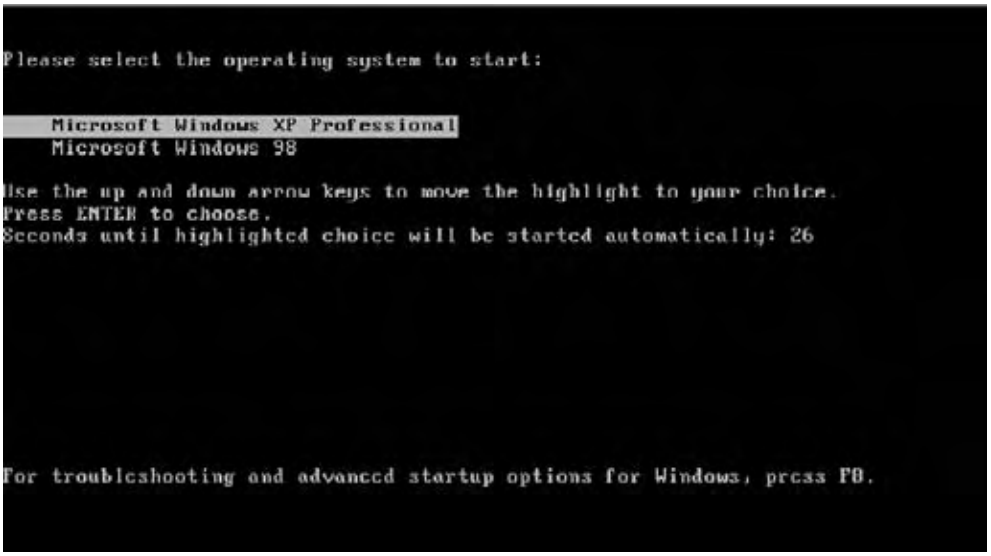
On disks that are configured to boot Windows XP, the volume boot record program looks for a program called NT Loader, or NTLDR for short. This program determines the file system of the default boot drive, which should be either FAT or NTFS.

If the operating system had been put to sleep or placed in hibernate mode, and the NTLDR finds a HIBERFIL.sys file, it will reload the contents of the file to RAM and resume operations at the point the system went to sleep.

Otherwise, the NTLDR looks for the BOOT.ini file in the root directory of the default drive.

NTLDR will display the versions of the operating system in a "boot menu" and waits a specified number of seconds before loading the first in the list (see Figure 12.2).

Figure 12.2 The Windows Boot Menu


Please select the operating system to start:
Microsoft Windows XP Professional
Microsoft Windows 98
Use the up and down arrow keys to move the highlight to your choice.
Press ENTER to choose.
Seconds until highlighted choice will be started automatically: 26
For troubleshooting and advanced startup options for Windows, press F8.

Once the operating system version has been selected, NTLDR runs the NTDETECT.com program to gather and verify the configuration of the computer.

After verifying the computer configuration the Hardware Abstraction Layer linked library HAL.dll is read into memory and then the Windows kernel NTOSKRNL.exe is loaded. Only after a series of driver checks and verifying that nothing has changed since the last run, or if the system was shut down improperly, does the NTLDR pass control to the kernel. If all goes well the end user sees the familiar login screen and is prompted to enter a username and password.

The number and types of checks that are run during this last phase of the boot process can be many depending on the configuration of the system, the version of the operating system, and the options that the administrator of the system has enabled. For more information on this exact process, see the Microsoft TechNet article *Troubleshooting the Startup Process* at http://technet.microsoft.com/en-us/library/bb457123.aspx.

Loading Linux

If the system to be loaded is a Linux-based system, then the volume boot record program will find the existence of either Grand Unified Boot Loader (GRUB) or the older Linux Loader (LILO). Both programs are boot loaders with LILO being the smaller and less feature-driven version for smaller systems. GRUB is the boot loader of choice on most modern systems since it allows for more interactive choices during the boot process.

LILO Booting

LILO is known as a Two Stage Loader, and stage one is stored in the 444 bytes of the MBR. Its primary functions are to gather the disk geometry information and load the second stage. The second stage is then loaded into memory and the Linux version is displayed on the screen. If there is more than one Linux kernel available in the LILO configuration file (/boot/boot.b) a menu of the available OS versions is displayed.

Either the menu choice or the default operating system version kernel and the contents of INITRD are loaded into RAM and LILO gives control to the kernel. Typically, LILO gives a command line login prompt only at completion.

The important files to LILO are:

- /sbin/lilo (the kernel installer)
- /boot/boot.b (the bootloader)
- /boot/map (tells LILO where the kernel is on the drive)
- /etc/lilo.conf (LILO configuration file)

GRUB Booting

GRB is *technically* also a two-stage loader as we will see. Like LILO, GRUB starts with the stage one loaded in the 444 bytes of the MBR. Stage one will load either stage 1.5 or 2 as required when the boot drive uses LBA mode or if the /boot partition is located beyond cylinder 1024. The stage 1.5 program code typically is found in the /boot partition.

The stage two boot loader is then read into memory and in turn loads the operating system kernel and the contents of INITRD. If multiple versions of the operating system are found in the configuration file GRUB will display a menu of those available and wait for a choice before loading that configured version of the OS. GRUB gives control of the system to the kernel and a login prompt either on the command line or via GUI window is presented to the end user.

The GRUB configuration file may be in several locations depending on the version of Linux, it can also be located on a removable USB drive! Some possible locations and names are:

- /boot/grub/menu.lst (SUSE and Debian)
- /boot/grub/grub.conf (Fedora uses this and the menu.lst above for compatibility.)

The Macintosh Boot Process

In this section, we will briefly examine the way an Apple Macintosh computer boots. The information here is for the Mac OS X version of their operating system using Intel based microprocessors. Older Motorola chipset Macintosh computers use a much different boot process.

OS X uses Open Firmware that is very much like the BIOS noted earlier. The Open Firmware that Apple uses in the Macintosh is based on the IEEE-1275 standard.

EFI and BIOS: Similar but Different

Just like any other computer on the market, when the power switch is activated on a Macintosh, the system goes through a Power On Self Test (POST), resets the microprocessor, and starts the execution of initialization code, which is the Open Firmware instead of BIOS.

Like the BIOS, Extensible Firmware Interface (EFI) checks the configuration of the machine and loads any device ROMs that it finds into memory. It then looks for a default boot device… and here is where it gets interesting. There are numerous optional startup functions that EFI can perform based on user input. Single keys, known as "snag keys," can be pressed that will allow the system to boot from specific devices.

- Pressing the C key will attempt to boot from the CD/DVD-ROM drive.
- Pressing the D key will attempt to boot from the first hard disk drive.
- Pressing the N key will attempt to boot from the Network Interface Controller (NIC).
- Pressing the Z key will attempt to boot from the ZIP drive.

It is also possible to enter the EFI interactive console mode by pressing the cmd-opt-O-F key combination during power up. (Note: If you are like me and just tried this before reading on, typing mac-boot at the prompt will let the Macintosh finish booting.) You should read a good source of Open Firmware/EFI commands before trying the console mode. An excellent mirror of the Open Firmware Working Group is at http://bananjr6000.apple.com/1275/.

The EFI program is located in the BOOT.efi file. This is the portion of the boot loading process that loads the OSX kernel and starts the user interface.

DARWIN

To many die-hard Macintosh users the move to OS X wasn't immediately seen as a move to the open source Unix environment. It wasn't long before they realized their beloved Mac was now a Unix machine. When you look at the roots of OS X, a large number of open source modules and programs were obtained from other groups including Carnegie Mellon, FreeBSD, GNU, Mach, Xfree86, NEXTSTEP, and OPENSTEP.

The OS X Kernel

In a nutshell the real OS X is when the combination of several components come together. XNU is the actual OS X kernel name on the boot drive. It is comprised of the following modules:

- Mach – Provides the service layer to the kernel
- BSD – Provides the primary system program interface
- I/O Toolkit – Provides driver support
- LIBSA & LIBKERN –Kernel libraries
- The Platform Expert – A motherboard-specific hardware abstraction layer
- Apple I/O components – The unique Mac interfaces

Apple uses proprietary components to invoke the Macintosh look and feel to the open source products listed. Carbon, Cocoa, Quartz, OpenGL, QuickTime, and the Aqua interfaces are just a few of the unique interfaces that make the Macintosh so special.

Notes from the Underground…

Bad Guy Won't Give You The Password? No Problem!

If you need to investigate a Macintosh that is running OS X and you need to access a program on a booted forensic copy of the subject's drive, and he won't give you his login password, don't worry. If you have any version of the Macintosh OS X boot CD or DVD, place that in the examination system and hold down the C key to boot from the CD/DVD drive.

When the system asks if you want to install/reinstall OS X, choose the Password Reset Utility from the drop-down menus at the top of the screen. You will be shown a list of users and you can pick one or all of them and change the password of the accounts to something you know. Problem solved!

Macintosh Forensic Software

Only recently has the Macintosh begun to be accepted in the forensic community. Listed next are just a few of the tools that can make forensics of OS X systems easier.

As with all forensic tools, the examiner should have a solid understanding of how tools work and should be able to prove by demonstration that each finding produced by the tool can be duplicated in a court of law.

BlackBag Forensic Suite

BlackBag Technologies, Inc. is one of the few providers of forensic software for the Macintosh platform. Its *Macintosh Forensic Suite* is a collection of 26 modules that can be launched individually or from the *Forensic Suite Toolbar* (see Figure 12.3).

Figure 12.3 The Forensic Suite Toolbar Is a Fast Way to Launch Programs in the Suite

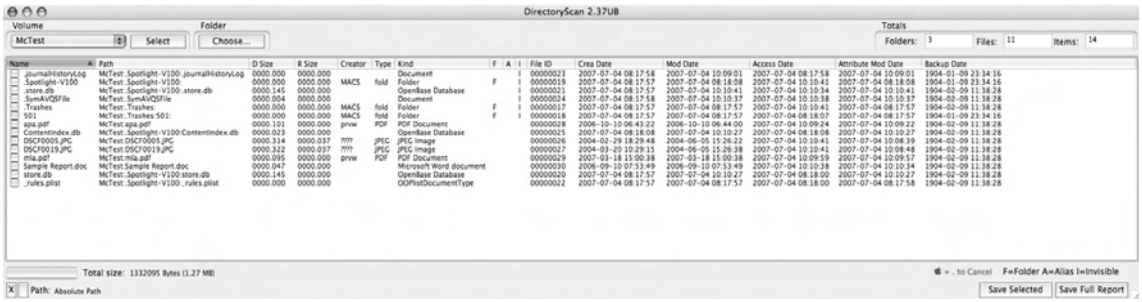


Directory Scan

The Directory Scan utility allows you to view all the files and folders on a Macintosh volume (see Figure 12.4). A volume can be any mounted storage device including USB or Firewire devices. All files, including invisible files, can be examined to include Data Fork/Resource Fork data sizes, Creator and Type codes, and all important date/time stamps.

You can select individual files and folders for export to a new directory for further examination as well as printing a comprehensive report on all the files viewed or selected in the main window.

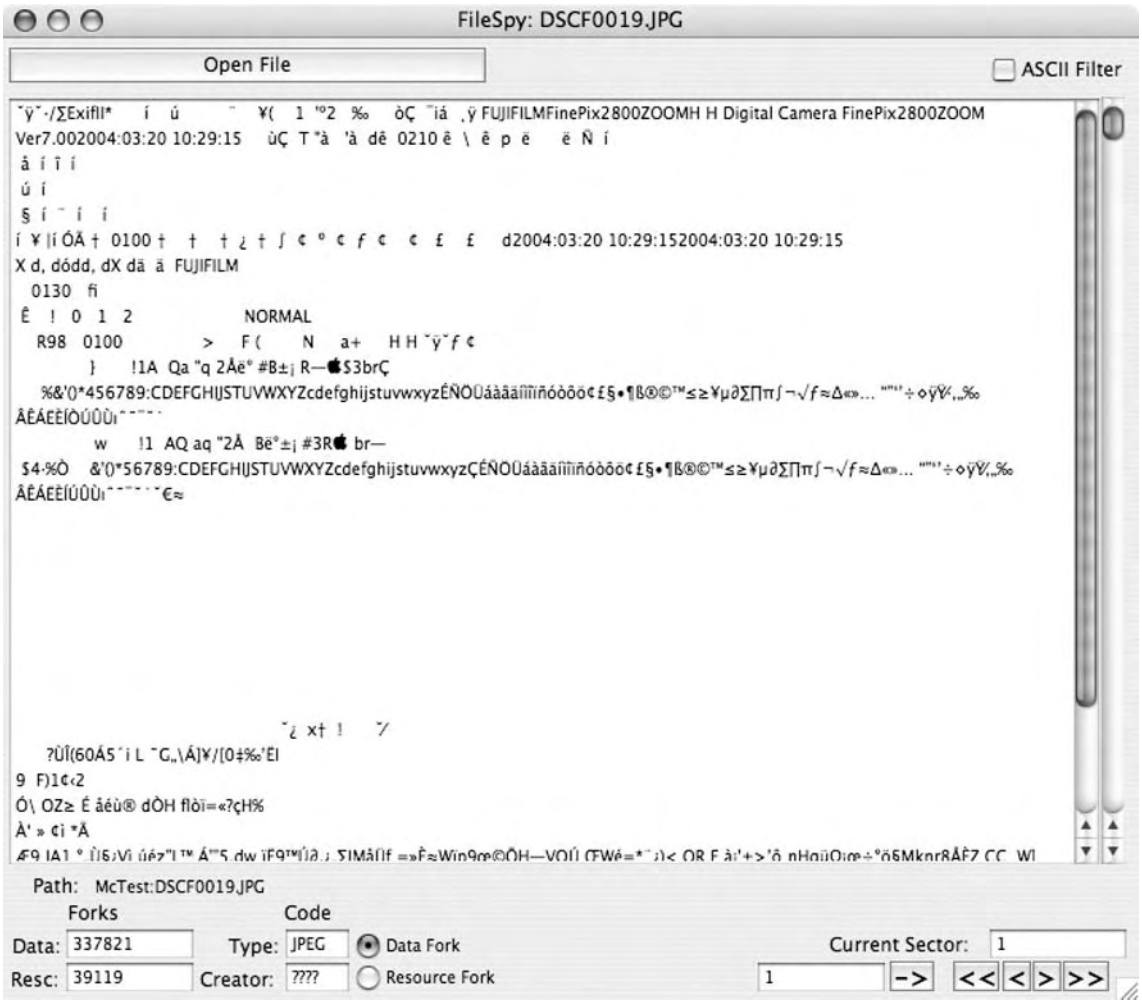
Figure 12.4 A Fully Expanded Directory Scan Window Can Be Quite Large



FileSpy

When you need to take a quick look inside of a file that has forks, FileSpy is a good tool (see Figure 12.5). This utility allow you to view either fork in a file, see the relative sizes of each fork, and move to any sector of a file directly. The utility even includes an ASCII filter to aid in file viewing.

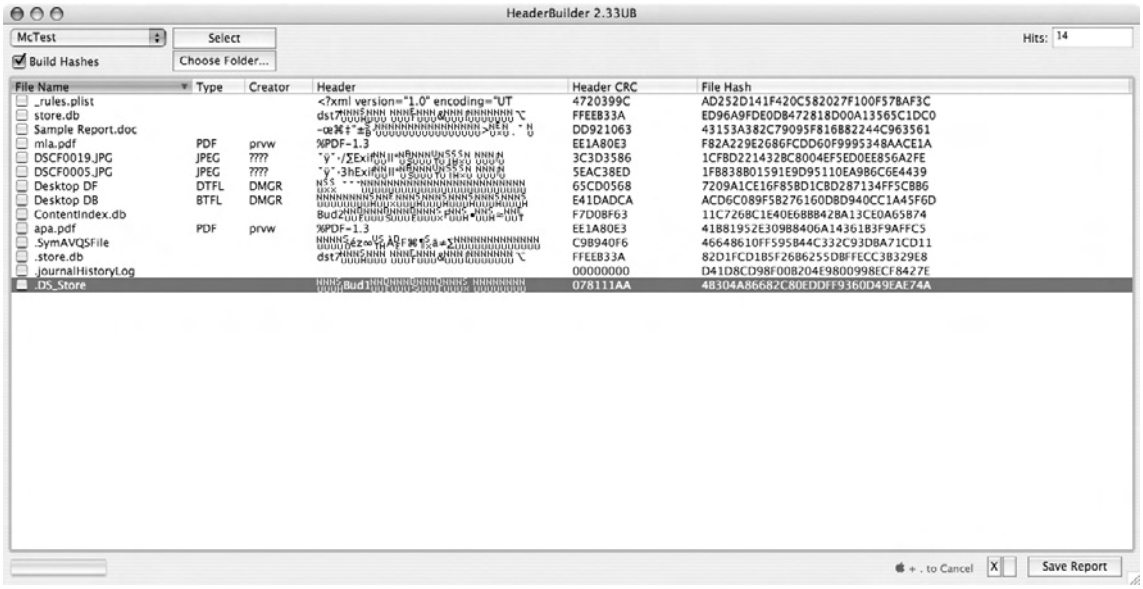
Figure 12.5 FileSpy Allows You to See the Raw EXIF Data in a JPG File



HeaderBuilder

Because the header is a calculated portion of Macintosh files, changing the header or repairing one can be time and math intensive using a traditional hex editor. HeaderBuilder makes this an easy task by allowing you to make the changes and then generate the CRC32 checksum and the MD5 hash of the file immediately (see Figure 12.6).

Figure 12.6 HeaderBuilder Makes Changing Headers Easy and Shows MD5 Hashes Quickly



Other Tools

Other utilities in the Forensic Suite include:

- Breakup – Splits large folders or files into more manageable sizes.
- Comment Hunter – Looks in the Comment fields of Mac files for keywords.
- DCFLDDassistant – Launches the Macintosh version of DCFLDD.
- File Searcher – Looks for specific filenames or Type/Creator codes.
- GraphicView – Uses the QuickTime engine to view files or movies.
- HFS Extractor – Converts image file formats (Sfaeback, Linux, DD, FWB).
- ImageBuster – Searches image files for keywords.
- ListBuilder – Allows you to create keyword lists in native languages (Spanish, Russian, etc.).
- LockMaster – Allows you to quickly lock or unlock a large number of files/folders.
- MacCarver – Lets you carve image files from within a container.
- PhantomSearch – Allows you to capture all the invisible files of a volume.

- Typer – A very fast little utility that shows/changes the Type/Creator for a given file.
- VolumeExplorer – HFS partition analyzer.

Carbon Copy Cloner

Mike Bombich has created a handy utility called Carbon Copy Cloner (CCC) for making backups or copies of important data on your Macintosh. It is a front-end for several less than intuitive utilities that are part of OS X.

As the name implies CCC can clone one hard disk to another when you use its default options. This copy can also be made to an image file on another drive, but it should be noted that this is not a forensic copy of the original (see Figure 12.7).

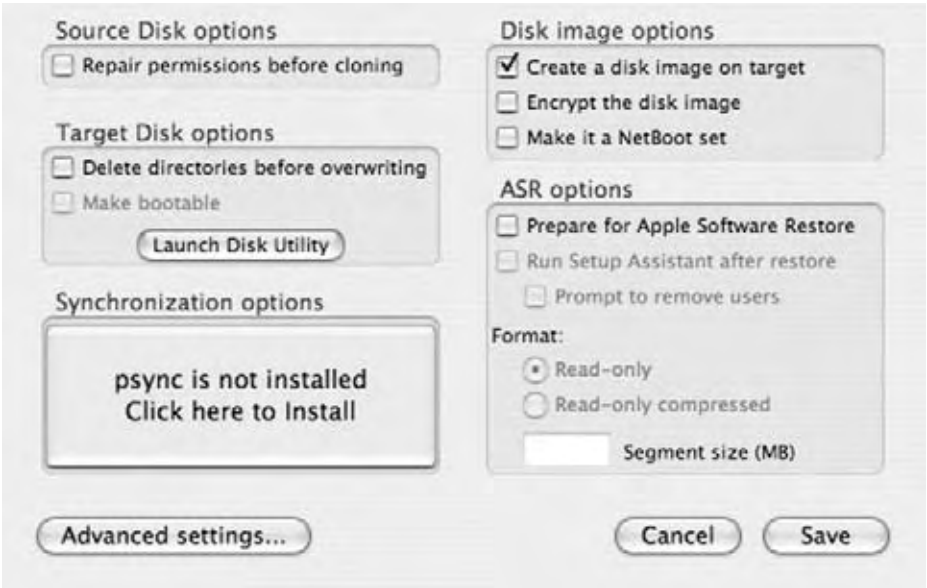
Documentation is available at the Bombich Software site: www.bombich.com/software/ccc.html.

Figure 12.7 Selecting the Source and Destination Drive Is a Simple Matter



Only Macintosh formatted volumes can be "cloned" using CCC; any other DOS or Unix formats are not recognized in the drop-down menus. If you do not have psync installed, you can install it from the Preferences menu (see Figure 12.8).

Figure 12.8 The Preferences Menu



Note the list of files that are not copied in the Advance Settings Dialog (see Figure 12.9).

Figure 12.9 The Advance Settings Dialog


Advanced Settings
Items in this list will be deleted from the target disk at the end of the clone. PLEASE read the "Advanced Features" section of CCC's Help documentation for more details on the items in this list.
Items to be deleted
Library/Caches/com.apple.LaunchServices.5B.csstore
Library/Caches/com.apple.LaunchServices.LocalCache.csst
Library/Preferences/SystemConfiguration/NetworkInterfac
System/Library/Extensions.kextcache
private/var/db/BootCache.playlist
private/var/db/NetworkInterfaces.xml
private/var/db/SystemConfiguration/com.apple.PowerMar
private/var/db/volinfo.database
private/var/vm/swapfile*
Reset to Defaults
Preflight script
Postflight script
Type the absolute path of the desired script in the fields above. Make sure the scripts are executable!
Save

The CCC documentation goes into more detail on the list of files that are not copied during the clone process. This can be because of PowerMac to Intel Macintosh copying causing problems. When in doubt check the reference material or the online forum.

MacDrive6/7

Technically Mediafour MacDrive 6 or 7 is not Macintosh software; it really is a Windows program that lets you mount and read Macintosh formatted volumes. MacDrive 6 is for Windows 2000 and 98SE; MacDrive 7 is for Windows XP (see Figure 12.10), Vista, and Server 2003.

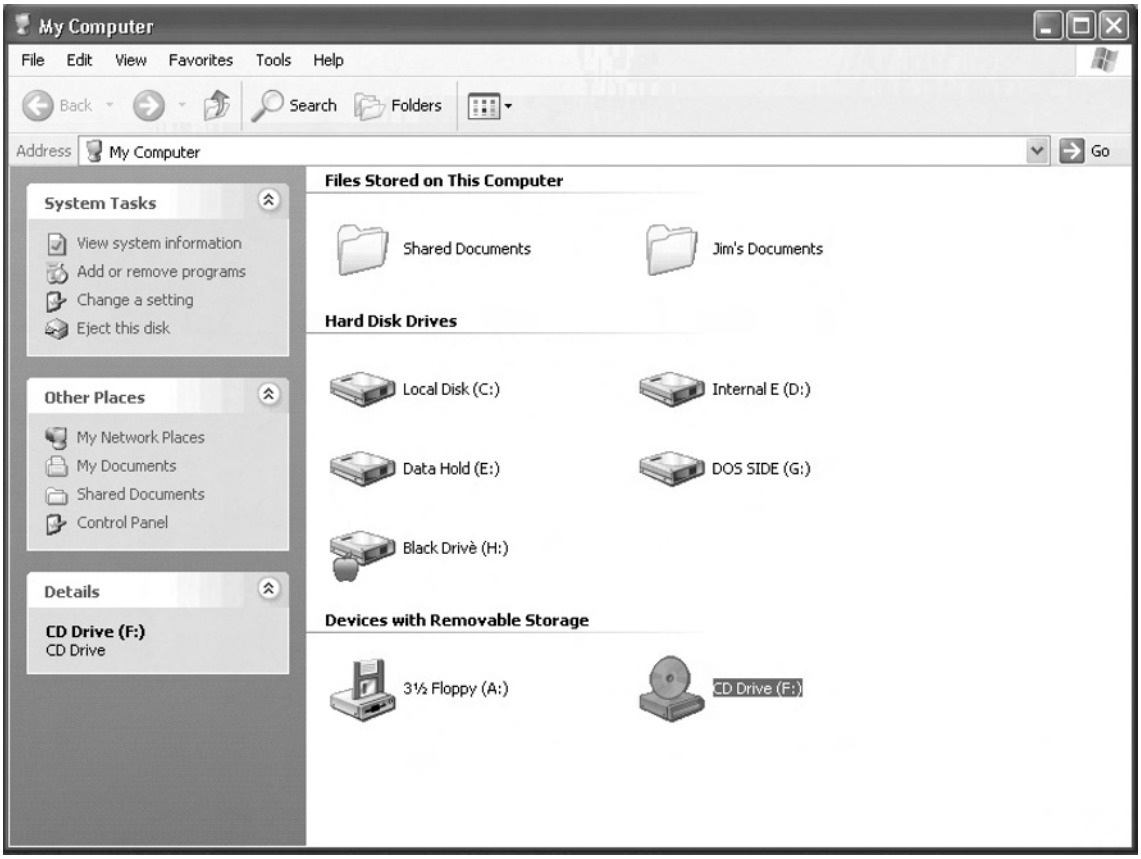
If you have a Windows-based workstation and need to quickly view some files on a Macintosh volume, this utility can be very helpful.

Figure 12.10 The Main MacDrive7 Screen (Windows XP version)



From the main menu you have the quick choices of Exploring a Macintosh volume, burning a Macintosh formatted CD or DVD, and formatting or repairing a Macintosh formatted volume. Mounted Macintosh volumes are shown with a small red Apple logo (see Figure 12.11).

Figure 12.11 Macintosh Options



Summary

Now you can see how the computer progresses through the steps of booting up. The power supply generates the Power-Good signal that resets the microprocessor and starts the loading of the BIOS from the Boot ROMs. Then the computer checks all its basic functions during the Power On Self Test (POST). If the POST is successful the computer then progresses through the different phases of loading the necessary drivers and initialization code that eventually leads to the computer finding the Boot Loader for a given operating system, which it then loads and runs. Finally the kernel of the operating system is loaded and the computer completes the boot process when the login prompt or command prompt is displayed and the computer is waiting for input.

You should also be familiar with the Master Boot Record (MBR) and how it is used in the boot process. The LILO and GRUB boot loaders for the Linux operating systems should be familiar now along with the Extensible Firmware Interface for the Macintosh OS.

Frequently Asked Questions

Q: Where is the Master Boot Record located?

A: The Master Boot Record is always located at Cylinder 0, Head 0, Sector 0 of a boot device or disk.

Q: How large is the Master Boot Record?

A: The Master Boot Record (MBR) is 512 bytes long and always ends with the end-of-file marker of h55AA.

Q: Do all Power On Self Test programs send error codes to port h80?

A: No. Some vendors of BIOS programs send error codes to other port numbers. This is typically a function of the card interface type. ISA uses port h80, EISA can use port h30, and so on. Most vendors have the error code ports listed on their web support sites.

Q: Is it true that the MSDOS command FDISK can refresh the Master Boot Record?

A: Yes! It is an undocumented feature of the FDISK command. FIDSK /MBR is the command.

Q: If I don't want the versions of Windows that are listed in the Boot menu, can I edit BOOT.ini manually to remove the ones I don't want?

A: The answer here is yes and no. It is possible to simply remove the lines for the Windows versions you don't want; however, they will still exist and take up room on the disk. It is better to follow Microsoft's guidance on removing these properly.

Q: If Macintosh OS X is for Intel-based processor systems, will it run on my PC?

A: There has been a lot of discussion on some of the hacker sites about a "hacked" version of Mac OS X running on regular Intel boxes. Using one of these modified versions is a violation of licensing with Apple Computers.

Chapter 13

Windows and Linux Forensics

Exam objectives in this chapter:

- **Windows Forensics**
- **Linux Forensics**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

In this chapter we introduce the reader to both Windows and Linux-based forensic tools. Although it is likely that you will have personal preferences, a thorough understanding of the strengths and weaknesses of each type of system will aid you in understanding both of them and also provide you with a wealth of additional tools.

By limiting the choices to a single operating system you are severely restricting your options. As such the best choice is to become familiar with the complete range of options available and this means gaining an understanding of both Windows and Linux.

Often, the most fruitful process for gathering evidence involves the use of both Windows and Linux-based tools. There are a number of Windows tools that enable the collection of data from live systems. Many Linux-based tools, on the other hand, provide a depth of analysis rarely found in any Windows-based tool. The combination of both Windows and Linux allows for the introduction of the strengths of both tool sets while removing many of the weaknesses.

Windows Forensics

The first section of this chapter is designed to introduce the reader to the forensic process under Windows. Even the most diehard Linux enthusiast will occasionally come across a live Windows system. Rather than shutting the system down, a large amount of volatile evidence may be collected on a live system. The section objectives include:

- Locating and gathering volatile evidence on a Windows host
- Investigating Windows file slack for evidence
- Interpreting the Windows registry and memory dump information
- Analyzing Internet trace data and events
- Investigating the system state backups

Where Can You Locate and Gather Evidence on a Windows Host?

The term "evidence location" refers to the process of investigating and gathering information of a forensic nature and particularly of legal importance. This evidence aids in the investigation of both criminal investigations and civil suits.

The Windows operating system contains a number of locations that can act as a rich source of evidence. Information gathering through investigating hidden files can be extremely helpful to any investigation.

Even file attributes and time stamps are valuable. Often, a perpetrator may attempt to change a file's attributes in order to either cover their tracks or hide important data that may be present in the system. Collating time stamps, for instance, can aid in reconstructing the actions taken by the suspect.

Some of the more important sources of electronic evidence on a Windows host include:

- Files
- Slack space
- Swap file
- Unallocated clusters
- Unused partitions
- Hidden partitions

How Can You Gather Volatile Evidence?

One of the most crucial aspects of digital forensics is one of the most often overlooked. This is the gathering of volatile data as evidence. When investigating a Windows-based operating system for possible evidence or information and facts relevant to the case, it is important to ensure that you have collected all relevant volatile data. Volatile data maintains current information about the system, the registry, cache, and memory. As such, if the system is powered down, valuable information is lost and may not be recovered. With nonvolatile memory, however, the data is not lost when the power is cycled.

Some of the most crucial areas to check for evidence within volatile data include: registers, cache, physical and virtual memory, network connections, running processes, and disk (for instance, the page file). Any external device associated with the system should also be considered and checked for evidence (floppy disk, tape, CD-ROM, and printers). Captured data must then be gathered and saved in external devices so that it may be safely removed and kept offline at another location.

RFC 3227 lists the order of volatility in a Windows-based system as:

- Registers, cache
- Routing table, ARP cache, process table, kernel statistics
- Memory
- Temporary file systems
- Disk
- Remote logging and monitoring data that is relevant to the system in question

- Physical configuration, network topology
- Archival media

The Features and Advantages of Windows Forensics Tools

It should be clear that with the majority of systems running Windows throughout the world, a large percentage of all incidents and thus forensic analysis will be conducted against Windows hosts. The ability to gather data on a live Windows system is incredibly valuable. Although offline imaging and searching through disk images is the standard fare of a computer forensic analyst, the value of volatile information should never be overlooked.

Further, there are times when it may not be feasible to take a system offline for imaging. In the case of a large e-commerce site (such as eBay) where any downtime is calculated in the millions of dollars, when critical infrastructure is a factor, or when a system must be maintained online, the ability to capture, image, and analyze a system while it is live can be of immense value.

The tools listed in this section not only provide a means of forensically capturing images of drives, but also enable the capturing and analysis of critical volatile data that would be lost if the system was powered down. In fact, these tools may be used to determine not only the nature of an incident but also whether an incident occurred. When something appears to be wrong, the cause is not always clear.

Helix Live on Windows

Helix runs on Windows as a standard application. In this mode it is used to collect evidence from active or live Windows systems. Remember: the state of a live system is in constant flux—it is constantly changing. Working on a live system will disturb the system state. Even if you do "nothing," you change the state of the system by running an application, changing the virtual memory, and so forth. Alternatively, turning off the system will result in the destruction of the same evidence; that is, the loss of potentially important forensic information. Tools such as Helix may be used to collect volatile information. There are also situations where the system cannot be turned off. Critical infrastructure, medical devices, or even financial markets systems often have to be maintained in a live state and thus imaged in that manner. To aid in this process, Helix presents a portable forensics environment which provides access to many Windows-based tools. Helix presents these tools and options together in a forensically sound manner.

Helix Live Response is really about the tools. The CD contains static binaries for Linux, Solaris, and Windows using GNU utilities and Cygwin tools. These tools include Sysinternal's tools, Garners Forensic Acquisition Utilities suite, Foundstone's open source

tools, the Windows Debugger, and the Windows Forensic Toolchest, amongst others. All these tools are accessible from the Helix.exe application.

The Helix.exe graphical application only operates in a live Windows environment. It has been tested and successfully runs on Windows 98SE, Windows NT4, Windows 2000, Windows 2003, Windows XP, and even Windows Vista. The Helix.exe application will vary slightly on each of these operating systems. Remember that as Windows is live, a number of DLL files are used by Helix and the operating system during this process.

The Windows Live function of Helix is basically a GUI interface (called the Incident Response screen) to a number of tools and a Windows command-line interface. The key benefits are that Helix will be your actions and maintain the integrity of the command-line by ensuring that built-in Windows tools (such as netstat or ipconfig) are not run from the compromised system itself, risking evidence corruption or data corruption. This is because Helix maintains known safe versions of these tools. It is common practice for an attacker to replace system tools (such as ipconfig).

The majority of Windows command-line tools that you will need are available and may be accessed from the Helix.exe application on the CD. These include:

- .cab extractors
- ipconfig
- kill
- netstat
- process explorers and much more

Some of the other valuable tools and applications on the CD that you should become familiar with include:

- An inclusive file attributes browser
- The Rootkit Revealer from Sysinternals
- A screen-capture application
- The PC Inspector data recovery tool

These are the tools that will be most useful. These tools are found in the Helix CD Windows module, which is also useful for a variety of ordinary network tasks. Of course, Helix is a forensics tool first and foremost; thus powerful tools for imaging disks for analysis are included (as well as tools to direct the image files to remote analysis systems).

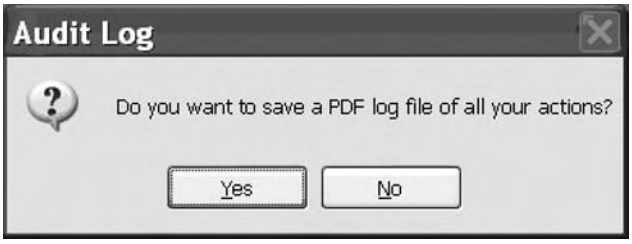
Damage & Defense…

Keep a Log of Your Actions!

Many forensic tools enable you to record your actions. The Helix Live CD allows you to save a log of what you did while using the program. Upon exiting Helix, you will be presented with the option to save your actions as a PDF log file (see Figure 13.1). Always save this file!

This log is your protection. Not only does it provide you with evidence of what you have done, saving on documentation, but it may also be used to validate your actions if they are disputed. Without this file it is just your word. Remember, logs are there for your protection too!

Figure 13.1 The Helix Live CD's Audit Log Option



The Tools Present in Helix CD for Windows Forensics

The following tools are all available from the Windows Live side of Helix. You should become familiar with their use.

Access PassView This application reveals the database password of password-protected MDB files created using Microsoft Access 95/97/2000/XP or with Jet Database Engine 3.0/4.0.

Astrick Logger This utility can reveal the passwords stored behind the asterisks in standard password text-boxes.

Drive Manager Drive Manager will identify drives which are of the same type. It displays the volume label as well as vendor information so that multiply CD/

DVD drives and removable drives such as USB thumb drives can be differentiated by their manufacturer's name, version, and revision date. The serial number can be seen as a unique ID for each drive.

FAU This is an Incident Response tool that can be used to image a system's memory as well as any attached devices.

Forensic Server Project The Forensic Server Project (FSP) is a proof of concept tool for retrieving volatile (and selected non-volatile) data from potentially compromised systems. The FSP consists of several Perl scripts and third-party utilities. The server component of the FSP is run on an investigator or administrator's system, and handles all data storage and activity logging. The client components of the FSP are on the CD so they may be run on the compromised system and connect to the remote server. Data is copied to the server component via TCP/IP.

FTK Imager FTK Imager can acquire physical device images and logically view data from FAT, NTFS, EXT 2 and 3 as well as HFS and HFS+ file systems. It also provides the ability to create multiple images from a single source and/or multiple images simultaneously. FTK Imager generates DD, SMART, and Encase® images and reads several other industry standard formats. Using Isobuster technology, FTK Imager provides ready access to CDFS and DVD file systems and allows access to include multi and open session CDs.

Galleta Galleta parses the information in a cookie file and output the results in a field-delimited manner so that it may be imported into a spreadsheet program.

HoverSnap HoverSnap is a snapshot tool with JPG, PNG, BMP, and GIF support. HoverSnap can take snapshots of the full screen, active window, or a selected area. It also captures layered windows (alpha-blended ones under 2K/XP). It enables the setup of a capture folder or filename and format. It will reduce the capture size. It allows an auto-generate filename option that will add the time stamp (date/time) to your filename so you can take several captures without having to change the filename.

IECookiesView IECookiesView is a small and handy utility that displays the details of all cookies that IE has stored on a computer. It will also sort the cookies, delete selected ones, view detailed information about each, and even save the cookies to a readable text file. Over a network, it can watch the cookies of other computers. If it has read permission on the cookies folders of other users it can view their cookies (admin rights). IECookiesView also allows the viewing of references to deleted cookies that are still stored in the index.dat file.

IEHistoryView IEHistoryView allows the viewing and modification of the history of visited websites in Internet Explorer. In addition it also may be used to export all or selected items to HTML reports, view detailed properties for selected

entries, and sort. The program allows you to access the history of other user accounts or network computers as well, provided that it is run with the proper access rights.

IRCR The Incident Response Collection Report is a script to call a collection of tools that gathers and/or analyzes data on a Microsoft Windows system. It provides a snapshot of the system. The tools are generally oriented towards data collection rather than analysis.

Mail PassView Mail PassView is a small password-recovery tool that reveals the passwords and other account details for Outlook Express, Microsoft Outlook 2000 (POP3 and SMTP Accounts only), Microsoft Outlook 2002/2003 (POP3, IMAP, HTTP, and SMTP Accounts), IncrediMail, Eudora, Netscape 6.x/7.x, Mozilla Thunderbird, Group Mail Free, and Yahoo! Mail email clients.

MEMDump The MEMDump utility is designed to dump or copy any part of 4 GB linear memory address space under MS-DOS and Windows 9x DOS to a console, text, or binary file.

MessenPass MessenPass allows the recovery of passwords from a wide variety of popular Instant Messenger programs, including MSN Messenger, Windows Messenger, Yahoo Messenger, ICQ Lite 4.x/2003, AOL Instant Messenger, AOL Instant Messenger/Netscape 7, Trillian, Miranda, and GAIM. When run, it will recover both the username and passwords. The list can be exported to HTML or saved as text file. MessenPass can only be used to recover the passwords for the current logged-on user on your local computer.

MozillaCookiesView MozillaCookiesView is an alternative to the standard "Cookie Manager" provided by Netscape and Mozilla browsers. It displays the details of all cookies stored inside the cookies file (cookies.txt) in one table, and allows you to save the cookies list into text, HTML, or XML file, delete unwanted cookies, and backup/restore the cookies file.

Network Password Recovery Network Password Recovery can retrieve all network passwords stored on a system for the current logged-on user; this includes any .NET Passport accounts that are stored locally.

Pasco Pasco parses the information contained within an index.dat file and outputs the results in a field delimited manner so that it may be imported into a spreadsheet program.

PC Inspector File Recovery PC Inspector File Recovery is a data recovery program that supports the FAT 12/16/32 and NTFS file systems. It can discover partitions automatically even where the boot sector or FAT has been erased or damaged. It recovers files with the original time and date stamp.

PC On/Off Time This application is a time tracking tool that shows the times a computer has been active during the last 3 weeks. It requires no previous setup.

Process Explorer The Process Explorer display consists of two sub-windows. The top window displays a list of the currently active processes, including the names of their owning accounts. The lower window displayed depends on the mode that Process Explorer is in. When the application is in handle mode you'll see the handles that the process selected in the top window opened. If it is in DLL mode you'll see the DLLs and memory-mapped files that the process has loaded. Process Explorer includes a powerful search capability that will quickly show which processes have particular handles opened or DLLs loaded.

Protected Storage PassView Protected Storage PassView is a small utility that reveals the passwords stored on a computer by Internet Explorer and Outlook Express. The passwords are revealed by reading the information from the Protected Storage. These include all email and web site passwords where you chose "remember password" (not cookie passwords) as well as auto-complete passwords. This utility only shows the passwords of the currently logged-on user and cannot reveal the passwords of other users.

PsTools Suite These tools allow the management of remote systems as well as the local one. PsList is a tool that lets you view detailed information about processes, and the suite is continually growing. The tools included in the PsTools suite:

- **PsExec** Execute processes remotely.
- **PsFile** Show files opened remotely.
- **PsGetSid** Display the SID of a computer or a user.
- **PsKill** Kill processes by name or process ID.
- **PsInfo** List information about a system.
- **PsList** List detailed information about processes.
- **PsLoggedOn** See who's logged on locally and via resource sharing.
- **PsLogList** Dump event log records.
- **PsPasswd** Change account passwords.
- **PsService** View and control services.
- **PsShutdown** Shut down and optionally reboot a computer.
- **PsSuspend** Suspend processes.

Pst Password Viewer Password encryption over a PST file is very fragile. For each password-protected PST file, there are many passwords that can open it.

PstPassword provides three different passwords for each password-protected PST file. All three passwords provided by PstPassword will open the PST file without problems.

PTFinder TFinder searches a memory dump of a system running Microsoft Windows for traces of processes and threads. Some functional checks are also applied.

PuTTY SSH Client PuTTY is a free implementation of Telnet and SSH for Win32, with an xterm terminal emulator.

Reg Adds, changes, and displays registry subkey information and values in registry entries, for the Local User.

RegScanner RegScanner is a small utility that allows you to scan the Registry, find the desired Registry values that match to the specified search criteria, and display them in a list. It can also export the found Registry values into a REG file that can be used in RegEdit.

ReSysInfo ReSysInfo is a system information viewer for Windows. The tool has 25 total information modules: BIOS information, CMOS, desktop, DirectX, drives, environment, fonts, keyboard, locale, machine and APM, mainboard, MCI, memory, mouse, multimedia, network, OpenGL, passwords, ports, printers and fax, processes, processor, video system, general information about Windows, and a summary. ReSysInfo also includes a Report Wizard, which can export the information to plain text, HTML, or XML format.

Rifiuti Rifiuti will parse the information in an INFO2 file and output the results in a field delimited manner so that it may be imported into a spreadsheet program.

Rootkit Revealer Rootkit Revealer is an advanced rootkit detection utility. Its output lists Registry and file system API discrepancies that may indicate the presence of a user-mode or kernel-mode rootkit. Rootkit Revealer successfully detects all persistent rootkits published at www.rootkit.com, including AFX, Vanquish, and HackerDefender.

SecReport This app is a small suite of two command-line tools. SecReport will collect security-related information from Windows-based systems. Delta will compare any two reports either from two systems or from the same system after some time. It produces XML-based reports that may be viewed with IE 6.0 browser, and automatically creates an MD5 hash file for reports.

WinAudit WinAudit is easy to use; no special knowledge is required to use the program. It is a self-contained single file that needs no installation or configuration. It reports on most aspects of computer inventory and configuration. Results are displayed in a web-page format that is categorized for ease of viewing and text searching.

Windows Forensic Toolchest Windows Forensic Toolchest (WFT) was written to provide for automated incident response on a Windows system and collect security-relevant information from the system. It is a forensically enhanced batch processing shell capable of running other security tools and producing HTML-based reports in a forensically sound manner. WFT is designed to produce output that is useful to the user, but is also appropriate for use in court proceedings. It provides extensive logging of all its actions along with computing the MD5 checksums along the way to ensure that its output is verifiable. The primary benefit of using WFT to perform incident responses is that it provides an easy method of scripting responses using a repeatable methodology for data collection.

MD5 Generators

An MD5 generator is a means of ensuring the integrity of a file, application, or file system. These applications create a cryptographic hash of the bit-wise information in the data. This hash is practically impossible to replicate. As a result, if you run the utility again at a later time and it returns the same value on the data, you may be assured that it remains unaltered.

This added layer of protection is essential in digital forensics. To maintain the value of the chain of custody and to ensure the admissibility of evidence, you need to ensure that the data has not been tampered with or altered. An MD5 generator can aid in achieving this end.

Chaos MD5 Chaos MD5 is a free MD5 generator for Windows. Input a file into it to generate an MD5 checksum for that file. That means that it generates a unique signature for each and every file. Chaos MD5 does not require installation. Just copy it to the hard drive or USB device to run. The MD5 checksum that is generated is used for file identification or integrity checks.

Secure Hash Signature Generator The Secure Hash Signature Generator creates signatures that are unique to the data stored on a disk drive. These signatures are used to verify intentional or accidental tampering with the drive image. The software has the ability to scan any secondary or slave drives attached to a computer running a Windows operating system. It does not write to the drive. It includes three different algorithms used to generate the signature: MD5 (128-bit signature), SHA1 (160-bit signature), and CRC (32-bit signature).

MD5 Generator: Mat-MD5 Mat-MD5 is designed to check the MD5 value for each file processed and compare it with other MD5 strings. It can process one or more files and add the result value to a list. You can add your MD5 value to compare by typing it or by copying it from an external file.

MD5 Checksum Verifier 2.1 MD5 Checksum Verifier is a file integrity checker which uses the MD5 algorithm. It can create checksums of files and verify their integrity. Operation involves two steps: make check file and verify check file.

Pslist

Pslist displays process, CPU, and memory information or thread statistics for all processes that are presently running on the system. The information listed for each process includes the time the process has executed, the amount of time the process has executed in kernel and user modes, and the amount of physical memory that the OS has assigned the process. Command-line switches allow the viewing of memory-oriented process information, thread statistics, or all three types of data. Here is an example of the syntax:

```
pslist [-?] [-d] [-m] [-x][-t][-s [n] [-r n]][\\computer [-u username]
[-p password]] [name | pid]
```

Some common parameters include:

- **-d** Shows statistics for all active threads on the system, grouping threads with their owning process.
- **-m** Shows memory-oriented information for each process, rather than the default of CPU-oriented information.
- **-x** Shows CPU, memory, and thread information for each of the processes specified.
- **-t** Shows the tree of processes.

This tool can be downloaded from www.microsoft.com/technet/sysinternals/default.mspx.

fport

fport displays all open TCP/IP and UDP ports and maps them to the application that owns the port. The application is similar to netstat except that it maps those ports to running processes with the PID, process name, and path. It is very useful in locating unknown open ports and their related applications. The switches may be utilized by making use of either a / or a - prior to the switch.

The command-line switches are as follows:

- **/?** Usage help
- **/p** Sort by port
- **/a** Sort by application
- **/i** Sort by pid
- **/ap** Sort by application path

This tool can be downloaded from www.foundstone.com.

Psloggedon

Psloggedon displays both locally logged on users and users logged on through resources for either the local or a remote computer. It ascertains which users are logged on by examining the keys under the HKEY_USERS key. For each key that has a name that is a user SID (Security IDentifier), Psloggedon finds the corresponding user name and displays that particular user name. Psloggedon utilizes NetSessionEnum API to ascertain who is logged onto a computer via resource shares. Here is an example of the syntax:

```
psloggedon [-?] [-l] [-x] [\\computername | username]
```

Some common parameters:

- **-?** Displays the supported options and the units of measurement used for output values.
- **-l** Shows only local logons instead of both local and network resource logons.

What Is File Slack? How Can You Investigate Windows File Slack?

A Windows disk cluster is a fixed length block of data which is used to store files. File size will never match the size of one or more clusters; this is where slack space occurs. File slack is the space that exists on a Windows disk (or other data device) between the end of the file and the end of the last cluster. It may contain data from the system's memory. It is possible to find user names, passwords, and other sensitive information associated with the system in slack. Larger cluster sizes relate to an increase in the volume of file slack leading to a waste in storage space. This is a vital source of evidence to the forensic investigator.

The Windows operating system provides for the difference between the end of the file and the end of a cluster by filling the leftover space with data from its memory buffers if there is no space available. This data, which is chosen randomly from the system's RAM is called RAM slack since it is obtained from the memory of the computer. RAM slack contains information related to files that are created, viewed, modified, or copied since the system was last booted. If the system has not been powered down for a period of time, then the data stored in file slack can come from past sessions.

Encase software provides a number of tools to conduct complex investigations with precision and effectiveness. Encase enables a methodical approach to digital forensic investigations. This approach allows the forensic analyst to manage large volumes of digital evidence and view computer drive contents, including files and artifacts such as operating system, file system, and deleted files or file fragments located in file slack or unallocated space.

How Do You Examine File Systems?

Examining the file system is an indispensable component of forensically analyzing a computer based on the Windows operating system. Simple commands such as *dir /O:d* (which sorts files in a directory based on the date order) play a significant role during the investigation of a Windows-based system. This simple command enables the forensic investigator to ascertain the time and date of the installation of the operating system. The service packs, patches, sub-directories, drivers, and other eventual files may be examined by using this command. When checking a file system, the analyst must give more regard to those files that have been updated most recently and develop a timeline of activities. Being able to sort files into time and date order makes the creation of a timeline a far simpler process.

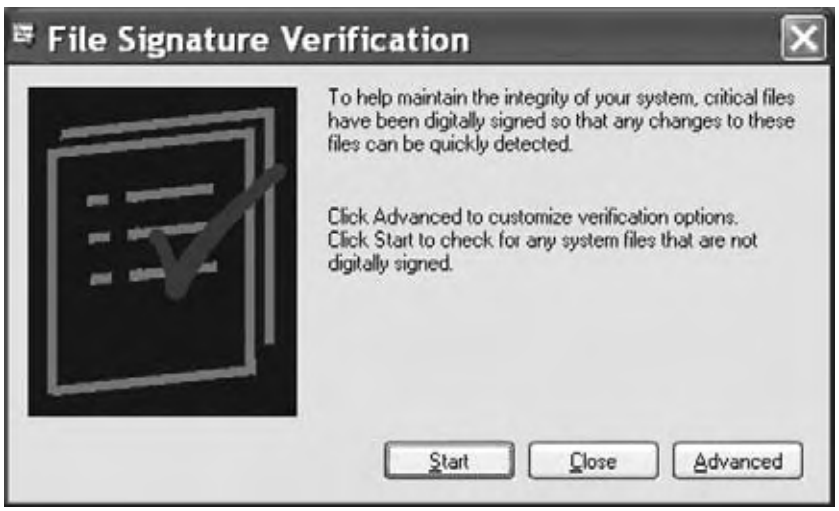
Built-in Tool: Sigverif

Sigverif.exe can analyze a system and report on any unsigned drivers found. To run the tool:

- Click **Start | Run**, type **Sigverif**, and click **OK**.
- Click the **Advanced** button (see Figure 13.2).
- Click **Look for other files that are not digitally signed**.
- Select the **WINNT | System32 | Drivers** folder, and then click **OK**.

After Sigverif.exe has completed the process of its checks, it displays a list of all unsigned drivers installed on the system. The list of all signed and unsigned drivers found by Sigverif.exe may be viewed in the Sigverif.txt file in the *Windir* folder, typically the *Winnt* or *Windows* folder. All unsigned drivers are marked as Unsigned.

Figure 13.2 Sigverif Is Included with Windows and Helps to Verify Digitally Signed Programs


File Signature Verification
To help maintain the integrity of your system, critical files have been digitally signed so that any changes to these files can be quickly detected.
Click Advanced to customize verification options. Click Start to check for any system files that are not digitally signed.
Start
Close
Advanced

The Word Extractor Forensic Tool

Word Extractor is an application that extracts human understandable interpretation from the computer binary format. It is similar to the UNIX command, *strings*. It offers users the following features:

- Replaces non-human words with spaces or dots for better visibility.
- Supports drag and drop and wrap text.
- Multitasks by responding to commands and processing large files.
- Saves results as TXT or RTF.
- Does not thwart registry or system with unwanted DLLs.

How Can You Interpret the Windows Registry and Memory Dump Information?

The Registry is an important component that should be examined during an investigation of a Windows system. Selected keys must always be checked and given more importance as they supply evidence for what services have been running, what services are presently running, and who had logged on to the system over a certain period of time. The primary keys that you will need to know for the exam include:

HKEY_LOCAL_MACHINE

- \Software\Microsoft\Windows\CurrentVersion\Run
- \Software\Microsoft\Windows\CurrentVersion\RunOnce
- \Software\Microsoft\Windows\CurrentVersion\RunOnceEx
- \Software\Microsoft\Windows\CurrentVersion\RunServices
- \Software\Microsoft\Windows\CurrentVersion\RunServicesOnce
- \Software\Microsoft\WindowsNT\CurrentVersion\Winlogon

The same keys must also be checked for the following:

- HKEY_CURRENT_USER
- HKEY_USERS.DEFAULT

Registry Viewer Tool: RegScanner

There are a variety of registry tools that either augment or supplant the Windows' Registry Editor Application (regedt32). Although it is simple to use, regedt32 does not come with a mass search function. It is not possible to find and display all instances of a given string in

the registry. The utility RegScanner from NirSoft provides this additional functionality. It provides a search function that can search for any value in the registry and display all the available instances of that value in a single view. From that view it allows a simple selection and jump function that takes you to the particular registry key for editing. It also provides a means to export results to a REG file so that they can be saved or loaded onto another computer.

RegScanner doesn't require installation. It may be unpacked from the distribution.zip file to any directory and run from there. When launched it provides an option to choose which base key to start a search from. It provides a dropdown menu with a number of common base keys. The search string can be case-sensitive or case-insensitive, and the match can be exact or within certain parameters. For instance, you can match against the data and the values, but not the key names themselves.

The application also provides a function to search for Unicode strings in binary values. RegScanner reports its results in a grid-type list, which can be saved as a REG file or exported to an HTML-format.

Microsoft Security ID

The Microsoft Security IDs are obtainable in the Windows registry key:

```
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\ProfileList
```

The ProfileList key holds the SIDs, on selecting individual ID value entry, and an associated user name is possible. There are some specific IDs enclosed under Microsoft Security that are viewed as follows:

- Encrypted File System Information
- Event Logging Utility
- Retrieve MAC Address
- Retrieve Security ID
- GUID to object
- Processor.vbs
- Registry Dump
- Registry Reference
- Registry Restoration
- Registry Scan

Summary of the Features and Importance of Memory Dump

Windows often creates a memory dump file when it crashes. This file contains valuable information that can aid in determining why the system stopped. For a memory dump to be produced in Windows the system requires a paging file that is at least 2 megabytes (MB) on the boot volume. Even if there is a large paging file on another disk but none on the boot volume, a memory dump will not function. Memory dumps are useful in aiding the diagnosis of system bugs and in analyzing memory contents during a program failure. They can contain information in binary, octal, or hexadecimal forms.

Where the system is running Microsoft Windows 2000 or later, a new file is originated every time the system stops abruptly. The Microsoft tool, dumpchk.exe, is designed to check memory dump files for information. The **systemroot | Minidump** folder contains a list of all small memory dump files. When conducting a forensic analysis of a Microsoft Windows system, always check for and obtain any memory dumps on the system. A memory dump file generally contains:

- The stop message and its parameters
- A list of loaded drivers
- The processor's context (PRCB) for the processes that stopped a regular normal operation of Windows
- The processor information and kernel context (EPROCESSES) for the processes that have been stopped
- The process information and kernel context (ETHREAD) for the thread that stopped
- The kernel-mode call stack for the thread that stopped the process from execution

Pagefile.sys and PMDump

Windows XP Professional uses paging file information to create a memory dump file in the systemroot directory. This dump file can be analyzed to supply some information as to the cause of a crash during offline analysis. It may also be analyzed using tools running on another computer. Post Mortem Dump or PMDump is a tool that performs the dumping of the memory contents related to any process to a file having its process stopped. This tool can be used for performing forensics analysis of a dump file.

What Is Virtual Memory?

Virtual memory is an imaginary memory location supported by the Windows operating system. It acts as an alternate set of memory addresses and expands the range of available memory. In this way slow hard drive memory is used to supplement physical RAM.

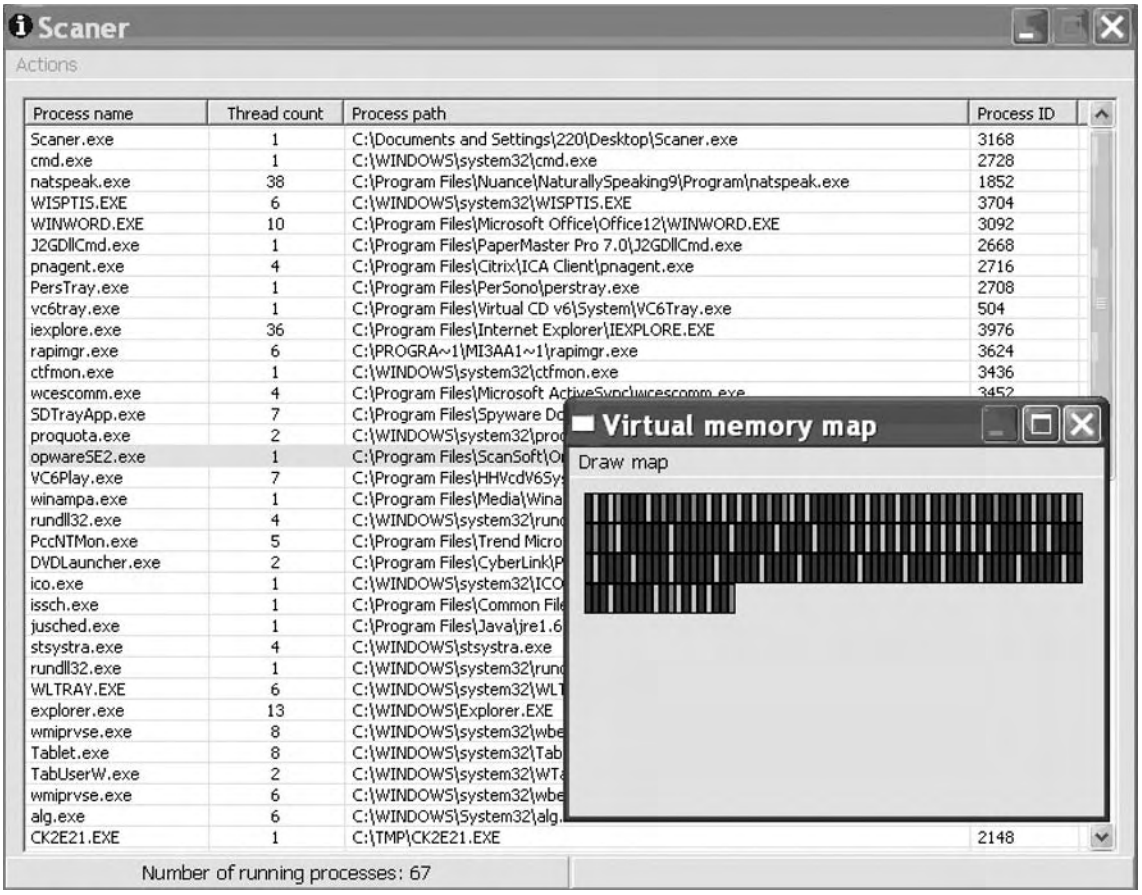
Programs use these virtual addresses for data that does not need to be addressed quite so often, to store instructions and data. When these memory address locations are called, virtual memory is converted into RAM.

In order to copy virtual memory into RAM, the operating system divides virtual memory into pages. Each page contains a fixed number of addresses that are stored on disk until called by the operating system. When pages are called by the operating system they are copied from disk memory to RAM, a process which translates virtual addresses into real addresses. This process is known as mapping. The process of copying virtual pages from disk to main memory is known as paging or swapping.

System Scanner

System scanner (available from www.codeproject.com/w2k/system_scaner.asp) is designed as a replacement to the Task manager.

Figure 13.3 System Scanner Provides a Visual Map of a System's Virtual Memory



The Windows task manager lacks the ability to fetch more specific info about the processes; that is supplied using the System Scanner (such as the IDs of all the threads, handles to DLLs, ability to suspend specific threads of a specific process, and, finally, an ability to view a process' virtual memory map—see Figure 13.3).

When the program is started, you are shown the main window, which shows all the currently running processes in your system, the number of threads per process, and the executable path. The status bar will show you the overall number of running processes. The list will be updated (by default) every 5000 milliseconds. Pressing **Enter** will force the update of the list.

From the main menu it is possible to force an update of the process list and enter **Options**, where you can customize the refresh time, turn hot track onoff (on by default) and customize the colors of the memory regions in the virtual memory map.

Right-clicking on any of the processes brings up a context menu which will uncover all the features of the program. This uncovers information about the process' threads, DLLs, virtual memory, and set process' priority, and offers a kill process function.

Integrated Windows Forensics Software: X-Ways Forensics and its Features

Based on WinHex, X-Ways Forensics is an advanced work environment for the digital forensic analyst. It provides a number of features such as:

- Forensically sound disk cloning and imaging
- Examining the complete directory structure inside raw image files, even if spanned over several segments
- Native support for FAT, NTFS, Ext2/3/4, CDFS, and UDF file systems
- Built-in interpretation of RAID 0 and RAID 5 systems and dynamic disks
- Viewing and dumping physical RAM and the virtual memory of running processes
- Various data recovery techniques and file carving
- File header signature database, based on flexible GREP notation
- Hard disk cleansing to produce forensically sterile media
- Gathering slack space, free space, inter-partition space, and generic text from drives and images
- File and directory catalog creation for all computer media
- Easy detection of and access to NTFS alternate data streams (ADS), even where other programs fail
- Mass hash calculation for files (CRC32, MD5, SHA-1, SHA-256)

- Manual NTFS decompression
- Support for the file systems HFS, HFS+, ReiserFS, Reiser4, UFS, UFS2
- Supported partitioning types: MBR, Windows dynamic volumes, GUID (GPT), Apple, unpartitioned (Floppy/Superfloppy)
- Gallery view for pictures
- Calendar view
- File preview, seamlessly integrated viewer component for 400+ file types
- Ability to examine e-mail extracted from Outlook (PST)*, Outlook Express (DBX), Mozilla (including Netscape and Thunderbird), generic mailbox (mbox, Berkeley, BSD, Unix), Eudora, PocoMail, Barca, Opera, Forte Agent, The Bat!, Pegasus, PMMail, FoxMail, *maildir* folders (local copies)
- Automated file signature check
- Ability to tag files and add them to customized report tables of notable items
- Automatic identification of encrypted MS Office and PDF documents
- Finds pictures embedded in documents (such as MS Word, PDF, and PowerPoint) automatically
- Internal viewer for Windows Registry files (all Windows versions); automated Registry report
- Viewer for Windows event log files
- Lists the contents of archives directly in the directory browser, even in a recursive view
- Logical search, in all or selected files/directories only, following fragmented cluster chains, in compressed files, optionally decoding text in PDF, WPD, and more
- Searches in both Unicode and various code pages
- Skin color detection (for example, a gallery view sorted by skin color percentage greatly accelerates a search for traces of child pornography and obscene images)
- Detection of host-protected areas (HPA), or ATA-protected areas

X-Ways Forensics also has the capability to write protect data to ensure authenticity and integrity. The product has a case management function integrated with automated activity logging (audit logs) and automated reports that can be imported and further processed by any other application that understands HTML, such as MS Word. It couples this with the ability to associate comments about files for inclusion in the report or for filtering.

How Can You Investigate Internet Traces?

Internet Explorer has several locations that should be checked for evidence. The following are the files and folders that should be investigated when analyzing Web browser activity:

- Cookies: In Windows 2000/XP, cookies can be found at the following location:
  1. <c:\Documents and settings\%username%\Cookies
  2. In Windows 95/98/ME, cookies can be found in the following folder: *C:\Windows\Cookies*
- History
- Typed URLs
- Temporary Internet files

Internet Explorer saves the record of all the sites visited by a user in a separate folder called the History folder. The URLs of all the websites visited by the user are present in this folder. There is another folder in the Documents and Settings folder, usually present in the C:\drive, which stores cookies visited by the user under the folder of each user. These cookies are generally stored in the Cookies folder. These files are present in the Temp folder. The location of temporary Internet files gives the name and number of sites visited by the user. The Temp files are important because they contain information about the user's activities in the various TMP files.

Traces Viewer

Traces Viewer as a tool that allows you to view all images, flash movies, pages, and other media files cached by the Internet Explorer browser. It also has a function that removes any web-traces made by Internet Explorer on a system. It doesn't wipe the evidence and it may be uncovered. It is available from www.bytescout.com/freeware.html.

IECookiesView

IECookiesView is a forensic utility that is used to display the details of all cookies that are stored by Internet Explorer. It provides the capability to sort cookies according to name and date. Additionally, it provides a search function that can find a cookie in the list by just specifying the name of a website. It will also copy cookie information onto the clipboard, and through remote login, can display cookies of other users of the same system or a different system.

IE History Viewer

Any URL that is typed into the address bar or any link that is followed in Internet Explorer automatically causes the address to be stored in the history index file. The IE History Viewer

parses all data from the history file, and prints the list of all visited URLs. This application provides the ability to save the list of addresses as a text, HTML, or XML file.

Cache Monitor

Cache monitor offers a real-time view of the present state of the systems cache. It can also verify the configuration of dynamic caches. This tool can prove helpful to the forensic analyst in verifying the cache policies and monitoring cache statistics. Few other tools provide the ability to monitor data flowing through cache and even the data present in the edge cache. The following cache statistics are available through Cache Monitor:

- **ESI Processors** Number of processes configured as edge caches.
- **Number of Edge Cached Entries** Number of entries currently cached on all edge servers and processes.
- **Cache Hits** Number of requests that match entries on edge servers.
- **Cache Misses by URL** Cache policy that do not exist on the edge server for the requested template.

How Do You Investigate System State Backups?

A system state backup is in effect a backup of the complete system. Everything that is present within the system will be copied as backup so that no data or information is lost when there is a system crash, a corruption of the driver files, or if certain system files stop the system from functioning properly. To perform a forensic analysis of evidence on a Windows system, backing up a system's registry is insufficient. An extensive backup of data is essential so that the system can be secured against any malfunctions.

A full system state backup saves the:

- Active Directory (NTDS)
- Windows Boot files
- COM+ class registration database
- Registry
- System volume (SYSVOL)
- The IIS metabase

The process to create a System State Backup is simple:

1. Go to Start | Programs | Accessories | System Tools | Backup
2. In the **Backup** tab, check the **System State** box.

3. Select the **Schedule Job** tab and select the **Add Job** button.
4. Select **Yes** and choose **media options***:*
    1. Media type
    2. Location
    3. Backup name
5. Select **Next** and check that the **Normal** option is selected, and then select **Next** again.
6. When backing-up to disk there is no need to verify data. Select **Next**.
7. Choose if you want to append to or replace an existing backup. Select **Next**.
8. Schedule the backup. Select **Next**.
9. Set the account with which the backup has the required permission to run.
10. Select **Finish**.

Investigating ADS Streams

A relatively unheard-of compatibility feature of NTFS is the Alternate Data Streams (ADS). These can provide an attacker with a method of hiding root kits or hacker tools on a compromised system which allows them to be executed without being detected by the systems administrator. Alternate Data Streams are strictly a feature of the NTFS file system. They may be used as a method of hiding executables or proprietary content.

Alternate Data Streams are simple to create and entail little or no skill to use. Common DOS commands such as **type** can be used to create Alternate Data Streams (ADS). These commands are used in conjunction with a redirect [>] and colon [:] to fork one file into another.

For example, the following command may be used to fork the common Windows calculator program with an ADS "malicious.exe":

type c:\malicious.exe > c:\winnt\system32\calc.exe:malicious.exe

Files with an ADS are almost impossible to detect using native file browsing techniques like command line or Windows Explorer. A third-party file checksum application can be used to effectively maintain the integrity of an NTFS partition against unauthorized Alternate Data Streams.

ADS Tool—LADS (List Alternate Data Streams):

- **Syntax** {file name}:{stream name}
- **Create** type textfile > visible.txt:hidden.txt
- **View** more < visible.txt:hidden.txt

Creating a CD-ROM Bootable for Windows XP

There are a number of tools such as the "Ultimate Boot CD-ROM" and "Bart's PE Builder" that are designed to assist you in creating a bootable Windows CD-ROM. These tools are appropriate for PC maintenance tasks and yield a complete Win32 environment with network support, a graphical user interface (800x600), and FAT/NTFS/CDFS file system support. All details of these tools are supplied below.

Bart PE (Bart Preinstalled Environment): Screenshot

Bart's PE Builder assists in building a "Bart Preinstalled Environment" (BartPE) bootable Windows CD-ROM or DVD from the original Windows XP or Windows Server 2003 installation/setup CD. This can restore any DOS boot disk. Supported Windows versions are:

- Windows XP Home Edition
- Windows XP Professional
- Windows Server 2003, Web Edition
- Windows Server 2003, Standard Edition
- Windows Server 2003, Enterprise Edition

Ultimate Boot CD-ROM

The Ultimate Boot CD-ROM allows floppy-based diagnostic tools to be executed from CD-ROM drives on Intel compatible machines. This does not require a separate operating system. This tool holds a number of diagnostic utilities that provide for sharing Internet access or surfing the web. Modifying NTFS volumes, restoring deleted files, scanning hard drives for viruses, and creating new NTFS volumes are possible with network support. This tool is available from www.ultimatebootcd.com/index.html.

The Ultimate Boot CD includes CPU Tests, memory tests, and peripherals tools.

CPU Tests

- CPU Burn-in
- Mersenne Prime Test
- StressCPU

Memory Tests

- Memtest86
- Memtest86+
- Windows Memory Diagnostic

- DocMem RAM Diagnostic
- TestMem4

Peripherals Tools

- Parallel port detection and test utilities
- ATAPI CD-ROM Identification
- CHZ Monitor-Test

The compilation contains tools in the following additional categories:

- CPU Information
- System Information
- Benchmarking Tools
- System Information
- Benchmark Tools
- BIOS Tools
- Hard Disk Installation Tools
- Hard Disk Diagnostic Tools
- Hard Disk Device Management Tools
- Hard Disk Wiping Tools
- Hard Disk Cloning Tools
- Hard Disk Low-Level Editing Tools
- Partition Tools
- Boot Managers
- File Tools
- Anti-Virus/Malware Tools
- Network Tools

The disk also supplies both DOS and Linux Boot Disks. A complete listing of the tools found on the Ultimate Boot CD-ROM is available from their website.

Linux Forensics

In this section of the chapter we will introduce the reader to computer forensics from the perspective of the Linux operating system. The objectives of the section are:

- Introduce Linux as a forensic tool
- Enable you to recognize the file systems and partitions in Linux
- Enable you to understand the challenges involved when using Linux
- Introduce you to the wide range of forensic tools that are freely available under Linux

Why Use Linux as a Forensic Tool?

With the wide range of Windows-based forensic tools, both in the open-source and commercial realm, it is easy for one to ask: Why learn Linux?

One of the key advantages of using Linux for forensics is the greater levels of control that it provides. Linux treats every device, from the hard drive to the monitor, as a file. One key benefit provided by Linux (and in fact UNIX in general) is its ability to mount drives in read-only mode. This capability means that Linux does not require a separate right blocker to ensure that data is not altered on the disk.

Another benefit of Linux is its flexibility. There are a wide range of Linux file systems and specialist forensic compilations designed to boot from a CD. Linux also has support for a large number of operating systems and is capable of mounting a large variety of file systems. Of particular benefit, the open source nature of Linux has resulted in the ability to access and image a large range of commercial file systems for little or no cost.

The availability of the Linux code coupled with this flexibility has resulted in a number of very powerful forensic distributions. Distributions such as F.I.R.E, Sleuth, Auditor, and Helix provide the forensic analyst with a wide range of choice coupled with a minimal capital outlay. These CDs provide file system support, platform support, recognize numerous file systems, and contain a large range of both command-line tools and GUI applications.

File System Description

Linux treats all devices as a file. As such, an entry is created in the file system for each hardware device recognized by the operating system. For instance, if a standard IDE hard drive is connected to the system, it will be listed under */dev/hdx*. These listings start from the letter "a" (or */dev/hda*) and increase in alphabetical order.

In a similar fashion, each partition of the drive is numbered from "1" on (starting with */dev/hda1*, for instance). To see a list of all the partitions that are available for a drive, type the following command: **fdisk –l /dev/hdx**.

The partitions are identified by their Linux terminology. A "*" indicates that it is a boot partition. The *if this* command also outputs the start and end cylinders for each of the partitions, as well as the number of blocks used by each partition. The output of the *fdisk* command also displays the file system type and partition identifier (ID).

The Primary Linux Directories

The majority of Linux distributions have a very similar directory structure. An example directory structure that will apply to most Linux systems and a description of the directory is supplied in Table 13.1.

Table 13.1 The Major Linux Directories and Their Functions

Linux directory path	Description of the Linux directory
/	The root directory is the lowest level directory in the Linux file system.
/bin	The bin directory contains executable programs which are needed in single user mode. These programs help boot the system or repair it.
/boot	The boot directory contains a static foals knitted by the boot loader. This directory should hold only files which are needed during the boot process.
/dev	This directory contains special or device files. These files refer to physical devices.
/etc	The *etc* directory contains consideration files which are local to the machine. Certain larger packages such as X11 have their own subdirectories with consideration files under */etc*.
/etc/opt	This directory contains host-specific configuration files for add-on applications which are installed in */opt*.
/etc/sgml	This directory is used to contain the configuration files for SGML and XML.
/etc/skel	The files contained in this directory are copied to a user home directory whenever a new account is created.
/etc/X11	This directory contains configuration files for the X11 windowing system.
/home	This directory is often configured as a root directory to place user home directories.
/lib	The */lib* directory holds shared libraries that are *ne* commands found in the root *Filesystem*.

Continued

Table 13.1 Continued

Linux directory path	Description of the Linux directory
/mnt	The */mnt* directory contains mount points to which file systems may be temporarily mounted.
/opt	The */opt* directory contains add-on packages in the form of static files.
/proc	The */proc* directory is a mount point for the proc file system. The proc pseudo-file system provides information about running processes and the Linux kernel. More detail is provided by proc(5).
/root	The */root* directory is commonly the home directory for the root user.
/sbin	The */sbin* directory holds system commands which are not usually executed by normal users.
/tmp	The */tmp* directory contains temporary files.
/usr	The */usr* directory generally contains local software, system libraries, games, and other applications. It is common for this directory to be mounted on a separate partition.
/usr/games	This directory commonly holds binary files associated with games.
/usr/include	Include files for the C compiler.
/usr/lib	This directory contains object libraries, including dynamic libraries, and some executables that are not generally invoked directly.
/usr/lib/X11	This directory contains data files associated with X programs, and configuration files for the X system.
/usr/local	Programs in consideration files which are local to the particular system are generally installed in or under this directory.
/usr/local/bin	This directory generally contains binaries for programs that are local to the host.
/usr/local/doc	This directory is for local documentation.
/usr/local/etc	This directory contains configuration files associated with the locally installed programs.
/usr/local/info	This directory contains the *info* pages associated with the locally installed programs.
/usr/local/man	This directory contains the *man* pages associated with the locally installed programs.

Table 13.1 Continued

Linux directory path	Description of the Linux directory
/usr/local/sbin	This directory contains binaries used for local system administration.
/usr/local/src	This directory contains the source code for locally installed software.
/usr/sbin	This directory contains program binaries for system administration. The binaries in this directory are not generally essential to the boot process and are not required for mounting */usr*. System files required for system repair should not be in this directory.
/usr/share/locale	This directory contains locale information.
/usr/share/man	The system's *man* pages are arranged in subdirectories according to their main page sections under this directory.
/usr/share/zoneinfo	Files containing data time-zone information are maintained under this directory.
/usr/src	This directory contains source files for the system.
/usr/src/linux	This directory was the traditional location to store the Linux kernel source.
/var	The *VAR* directory contains variable files which are expected to change in size. This includes spool and log files.
/var/lib	This directory maintains variable state information for the programs running on the host.
/var/local	District ridges to hold variable data files and programs located under */usr/local*.
/var/log	The standard log directory.
/var/mail	Location in many systems of user mailboxes.
/var/run	This directory contains run-time variable files such as files holding process identifiers (PIDs). Logged user information (such as *utmp*) is also maintained in this directory. Most file systems clear this directory on a reboot.
/var/spool	The spool directory contains queued files; *at*, *cron*, *lpd*, and *mqueue* will all create subdirectories to spool jobs underneath this directory.
/var/tmp	This directory is similar to */tmp* and stores temporary files for an unspecified length of time.
/var/yp	This directory contains database files for NIS.

Mount Command

Before a file system can be used under Linux it needs to be mounted (this statement also applies to the Mac and other UNIX variants). Linux will mount any file systems defined in */etc/fstab* at each system boot. It is important to note however that Linux will write to devices even when they are not mounted.

A common location to situate mount points is in the directory */mnt*. Only the *superuser* (or *root*) account can create directories in */mnt*. Redhat and several other flavors of Linux come with mount points already set up for a floppy, a CD-ROM, and a parallel port Zip drive. To export a file or directory from a server over a network, it may need to be explicitly specified by root account in the */etc/fstab* file.

Table 13.2 lists common mount points and their options.

Table 13.2 Common Settings Used with the Mount Command

Node:Mount_dir	Mount Point	File Type	Mount Options
/dev/fd0	/mnt/floppy	auto	noauto 0 1
/dev/cdrom	/mnt/cdrom	iso9660	ro, noauto 0 1
/dev/zip	/mnt/zip	Vfat	auto, exec 0 1
server:/exported	/mnt/server	Nfs	Defaults

Linux will allow you to analyze file systems from several different operating systems. All that is required is to mount the file system with the proper type and options. For example, using Sun UFS:

```
# mount -r -t ufs -o ro,ufstype=sun /dev/hdc2 /mnt
```

Always use the *noexec* and *nodev* options to prevent accidentally running programs from the partition being analyzed, and to ignore any device files it may contain. This is an example of a typical mount command:

```
# mount -o ro,noexec,nodev /dev/hdc1 /t
```

The Linux kernel can support a large number of file system types (the following information is taken from the command *man mount*):

> The file system types which are currently supported are listed in linux/fs/filesystems.c: adfs, affs, autofs, coda, coherent, devpts, efs, ext, ext2, hfs, hpfs, iso9660, minix, msdos, ncpfs, nfs, ntfs, proc, qnx4, romfs, smbfs, sysv, udf, ufs, umsdos, vfat, xenix, xiafs. Note that coherent, sysv and xenix are equivalent and that xenix and coherent will be removed at some point in the future—use sysv instead.

For this reason Linux makes an excellent forensic analysis platform. The ability to use a single set of utilities to access file contents and i-node information from a large number of system types for little cost is a key feature of all Linux forensic applications. Another benefit of Linux is the inclusion of "loopback" devices. Loopback devices provide the ability to mount a file containing an image copy (such as that obtained using *dd*) with your forensic analysis host.

Loopback devices allow the use of a file system within a file. This method is used for bootdisks, burning CD-ROMS, and for encrypted file systems. Further information may be found in the following commands and Linux FAQs:

- man losetup
- man mount
- Laptop-HOWTO
- Bootdisk-HOWTO
- Loopback-Encrypted-Filesystem-HOWTO

The Linux Boot Sequence

Linux starts the boot process by loading the kernel. Generally, the Linux kernel image will be contained in the boot directory (*/boot*). Either LILO or GRUB may be queried by viewing their consideration files (*/etc/lilo.conf* or */etc/grub.conf*) using the command **less**. The boot loader specifies the location of the "root device," which is in effect the boot drive and the kernel version which is to be booted.

Next, the Linux boot process starts initialization. In this stage *run level* and *startup* scripts are initialized and the terminal process is controlled. Initialization is controlled through the file */etc/inittab*. The *file/sbin/init* starts the initialization process. In Linux (like other UNIX variants), *runlevel* is used as a description of the system state. In Red Hat Linux some of the startup scripts of interest are:

- **/etc/rc.d/rc.sysinit** This script controls system initialization, file system mount, and checks, amongst other things.
- **/etc/rc.d/rc.*** For each value of * (0-6), the *runlevel* passed as an argument by the *init* process. This script calls the specified script for the *runlevel* as it starts.
- **/etc/rc.d/rc.local** This script is called from within the specific *runlevel* scripts. The script, *rc.local*, is a general purpose script that is modified to include commands needed at bootup and to customize a local server.

One of the more common shells deployed is the Linux BASH shell (Bourne Again Shell). Command shell can be thought of as equivalent to the DOS command.com shell. Some other common shells include:

- **Bourne** The traditional Unix shell.
- **ksh Korn** The de facto standard shell for Unix systems.
- **csh** The original C Shell.
- **tcsh** Extended version of the C Shell.
- **zsh** A powerful shell that integrates much of the functionality of *ksh*, *csh*, and *bash*.

The Challenges in Disk Forensics with Linux

There are a number of challenges that will be faced when Linux is used as a forensic tool to analyze a disk image. Some of these include:

- Linux is unable to identify the last sector on hard drives with an odd number of sectors.
- The majority of Linux tools either do not have a GUI interface or work better at the command line, which makes them difficult to use and increases the learning curve.
- Devices can be written to at all times. Although it is possible to mount an image as a read-only file system which will not be changed, it is still possible to write data directly to the physical drive using Linux.
- Although both commercial and open source software has bugs, Linux is often perceived as being riskier. This can influence the credibility of the tools.

Notes from the Underground…

Imaging Odd-Sized Drives with Linux

Although it is rare to find a tribe of an odd number of sectors, they could pose a problem for Linux. Linux cannot see the last sector of a drive with an odd number of sectors. As a result, a small amount of information may be lost in this situation.

It is recommended that either a BSD variant, commercial UNIX, or commercial Windows tool be used in this instance. Although it is unlikely that a single sector at the end of a disk will contain evidence that could sway a case, you never know.

Popular Linux Forensics Tools

Linux includes several simple utilities to make imaging and basic analysis of suspect disks and drives simpler. These include:

- **dd** Command used to copy from an input file or device to an output file or device.
- **dcfldd** A more advanced variant of *dd* that shows the progress while imaging.
- **sfdisk/fdisk** Used to determine the structure of a disk.
- **grep** Search files for instances of an expression or pattern.
- **loopback device** Allows user to mount an image without having to rewrite the image to a disk.
- **md5sum** and **sha1sum** Create and store an MD5 or SHA hash of a file or list of files (including devices).
- **file** Reads file header information in an attempt to ascertain its type, regardless of name or extension.
- **xxd** Command line hexdump tool.
- **ghex** and **khexedit** The Gnome and KDE (X Window interfaces) hex editors.

The Sleuth Kit

The Sleuth Kit (TSK) is a collection of command line tools based on The Coroner's Toolkit (TCT). The Autopsy Browser is a graphical interface to TSK. www.sleuthkit.org is the official website for The Sleuth Kit and Autopsy Browser. Both are open source digital forensic tools that run on UNIX systems (including Linux, OS X, FreeBSD, OpenBSD, and Solaris). These tools can be used to analyze NTFS, FAT, Ext2, Ext3, UFS1, and UFS2 file systems and several volume system types. The Sleuth Kit was developed by Brian Carrier.

The Sleuth Kit (formerly known as TASK) is a collection of UNIX-based command line file and volume system forensic analysis tools. The file system tools allow the examination of file systems associated with a suspect computer in a non-intrusive fashion. The tools do not rely on the operating system to process the file systems which means that deleted and hidden content is shown.

The volume system (media management) tools allow the examination of the disk layout and other media. The Sleuth Kit maintain support for DOS partitions, BSD partitions (disk labels), Mac partitions, Sun slices (Volume Table of Contents), and GPT disks. These tools enable the identification of a partition's location and can also extract partitions for analysis with file system analysis tools.

The Autopsy Forensic Browser is a graphical interface to the tools in The Sleuth Kit, which interfaces with the command line tools to simplify the process. Additionally, Autopsy provides case management, image integrity, keyword searching, and other automated operations that are not available in the individual tools themselves.

Tools Present in "The Sleuth Kit"

The tools in The Sleuth Kit take a disk or file system image as input. All of the file system tools support NTFS, FAT, Ext2/3, and UFS1/2 file systems. An up-to-date list is maintained at the TSK site: www.sleuthkit.org/sleuthkit/tools.php.

File system layer tools process general file system data, such as the layout, allocation structures, and boot blocks. The tool *fsstat* shows file system details and statistics including layout, sizes, and labels.

File name layer tools process the file name structures, which are typically located in the parent directory.

- **ffind** Finds allocated and unallocated file names that point to a given meta-data structure.
- **fls** Lists allocated and deleted file names in a directory.

Metadata layer tools process the metadata structures, which store the details about a file. Examples of this structure include directory entries in FAT, MFT entries in NTFS, and inodes in ExtX and UFS.

- **icat** Extracts the data units of a file, which is specified by its meta-data address (instead of the file name).
- **ifind** Finds the meta-data structure that has a given file name pointing to it or the meta-data structure that points to a given data unit.
- **ils** Lists the meta-data structures and their contents in a pipe delimited format.
- **istat** Displays the statistics and details about a given meta-data structure in an easy to read format.

Data unit layer tools process the data units where file content is stored. Examples of this layer include clusters in FAT and NTFS, and blocks and fragments in ExtX and UFS.

- **dcat** Extracts the contents of a given data unit.
- **dls** Lists the details about data units and can extract the unallocated space of the file system.
- **dstat** Displays the statistics about a given data unit in an easy to read format.
- **dcalc** Calculates where data in the unallocated space image (from dls) exists in the original image. This is used when evidence is found in unallocated space.

File system journal tools process the journal that is implemented in selected file systems. A journal records the metadata (and occasionally content) updates that are made. This may aid in the recovery of recently deleted data. Examples of file systems with journals include Ext3 and NTFS.

- **jcat** Display the contents of a specific journal block.
- **jls** List the entries in the file system journal.

Medium management tools take a disk (or other media) image as input and analyze its partition structures. Examples include DOS partitions, BSD disk labels, and the Sun Volume Table of Contents (VTOC). These can be used to find hidden data between partitions and to identify the file system offset for The Sleuth Kit tools. The media management tools support DOS partitions, BSD disk labels, Sun VTOC, and Mac partitions.

The tool *mmls* displays the layout of a disk, including the unallocated spaces. The output identifies the type of partition and its length, which makes it easy to use *dd* to extract the partitions. The output is sorted based on the starting sector so it is easy to identify gaps in the layout.

Image file tools work with the format of the image file. For instance, they can work with an image format that is a split image or a compressed image.

- **img_stat** This tool will show the details of the image format.
- **img_cat** This tool will show the raw contents of an image file.

Disk tools may be used to detect and remove a Host Protected Area (HPA) in an ATA disk. A HPA could be used to hide data so that it would not be copied during an acquisition. These tools are only available for Linux at the present time.

- **disk_sreset** This tool will temporarily remove a HPA if one exists. After the disk is reset, the HPA will return.
- **disk_stat** This tool will show if a HPA exists.

A number of other miscellaneous tools can also be found in the toolkit.

- **hfind** Uses a binary sort algorithm to look up hashes in the NIST, NSRL, Hashkeeper, and custom hash databases created by md5sum.
- **mactime** Takes input from the *fls* and *ils* tools to create a timeline of file activity.
- **sorter** Sorts files based on their file type and performs extension checking and hash database look-ups.
- **sigfind** Searches for a binary value at a given offset. Useful for recovering lost data structures.

Autopsy

The Autopsy Forensic Browser is an HTML front-end for The Sleuth Kit. As such, it provides a graphical interface to the command line digital forensic analysis tools in The Sleuth Kit. Together, they can analyze Windows and UNIX disks and file systems (NTFS, FAT, UFS1/2, Ext2/3). The Sleuth Kit and Autopsy are both Open Source and run on Linux/UNIX platforms.

The code of Autopsy is open source and all files that it uses are in a raw format. All configuration files are in ASCII text and cases are organized by directories. This makes it easy to export the data and archive it. It also does not restrict the use of other tools that may solve the particular problem better.

The Evidence Analysis Techniques in Autopsy

This section introduces the primary modes and functions of the Autopsy Forensic Browser. This collection of tools acts as a graphical front end to the Sleuth Kit and other related tools in order to provide the capabilities of analysis, search, and case management in a simple but comprehensive package.

A dead analysis occurs when a dedicated analysis system is used to examine the data from a suspect system. When this occurs, Autopsy and The Sleuth Kit are run in a trusted environment, typically in a lab. Autopsy and TSK provide support for raw, Expert Witness, and AFF file formats.

A live analysis occurs when the suspect system is being analyzed while it is running. In this case, Autopsy and The Sleuth Kit are run from a CD in a non-trusted environment. This is frequently used during incident response while the incident is being confirmed. Following confirmation, the system is acquired and a dead analysis performed.

The Autopsy Browser provides the following evidence search functionality:

- **File Listing** Analyze the files and directories, including the names of deleted files and files with Unicode-based names.
- **File Content** The contents of files can be viewed in raw, hex, or the ASCII strings can be extracted. When data is interpreted, Autopsy sanitizes it to prevent damage to the local analysis system. Autopsy does not use any client-side scripting languages.
- **Hash Databases** Look up unknown files in a hash database to quickly identify it as good or bad. Autopsy uses the NIST National Software Reference Library (NSRL) and user-created databases of known good and known bad files.
- **File Type Sorting** Sort the files based on their internal signatures to identify files of a known type. Autopsy can also extract only graphic images (including thumbnails). The extension of the file will also be compared to the file type to identify files that may have had their extension changed to hide them.

- **Timeline of File Activity** A timeline of file activity can help identify areas of a file system that may contain evidence. Autopsy can create timelines that contain entries for the Modified, Access, and Change (MAC) times of both allocated and unallocated files.
- **Keyword Search** Keyword searches of the file system image can be performed using ASCII strings and *grep* regular expressions. Searches can be performed on either the full file system image or just the unallocated space. An index file can be created for faster searches. Strings that are frequently searched for can be easily configured into Autopsy for automated searching.
- **Meta Data Analysis** Meta-data structures contain the details about files and directories. Autopsy allows you to view the details of any meta-data structure in the file system. This is useful for recovering deleted content. Autopsy will search the directories to identify the full path of the file that has allocated the structure.
- **Data Unit Analysis** Data Units are where the file content is stored. Autopsy allows you to view the contents of any data unit in a variety of formats including ASCII, hexdump, and strings. The file type is also given and Autopsy will search the meta-data structures to identify which has allocated the data unit.
- **Image Details** File system details can be viewed, including on-disk layout and times of activity. This mode provides information that is useful during data recovery.

Autopsy provides a number of functions that aid in case management. In particular, investigations started within Autopsy are organized by cases, which can contain one or more hosts. Each host is configured to have its own time zone setting and clock skew so that the times shown are the same as the original user would have seen. Each host can contain one or more file system images to analyze. The following functions within Autopsy are specifically designed to aid in case management:

- **Event Sequencer** Time-based events can be added from file activity or IDS and firewall logs. Autopsy sorts the events so that the sequence of incidents associated with an event can be easily determined.
- **Notes** Notes can be saved on a per-host and per-investigator basis. These allow the investigator to make quick notes about files and structures. The original location can be easily recalled with the click of a button when the notes are later reviewed. All notes are stored in an ASCII file.
- **Image Integrity** Since one of the most crucial aspects of a forensics investigation involves ensuring that data is not modified during analysis, Autopsy will generate an MD5 value for all files that are imported or created by default. The integrity of any file that Autopsy uses can be validated at any time.

- **Reports** Autopsy can create ASCII reports for files and other file system structures. This enables investigators to promptly make consistent data sheets during the course of the investigation.
- **Logging** Audit logs are created on a case, host, and investigator level so that all actions can be easily retrieved. The entire Sleuth Kit commands are logged exactly as they are executed on the system.

SMART for Linux

SMART by ASR Data is a commercial GUI forensic tool for Linux. It is a software utility that has been designed and optimized to support data forensic practitioners and Information Security personnel in pursuit of their respective duties and goals. The SMART software and methodology have been developed with the intention of integrating technical, legal, and end-user requirements into a complete package that enables the user to perform their job most effectively and efficiently.

SMART includes a number of features that allow it to be used in a wide range of situations, including:

- "Knock-and-talk" inquiries and investigations
- On-site or remote preview of a target system
- Postmortem analysis of a dead system
- Testing and verification of other forensic programs
- Conversion of proprietary "evidence file" formats
- Baselining of a system

SMART lists the connected storage devices in its main window. Devices are listed based on the hardware connection type. Both drive partitions and unpartitioned space are listed as sub-items in the same layout as is on the device. Information such as the device make and model, storage capacity, and even partition disk usage are displayed together on a simple interface. Detailed information, such as a serial number or raw byte count, may be accessed by right clicking on the device.

Penguin Sleuth

The Penguin Sleuth Bootable CD is a forensic tool collection on CD by Ernest Bacca and is available from www.linux-forensics.com. Penguin Sleuth contains many useful tools. It is an adapted version of the Knoppix Linux Live CD that has been modified to include tools that are useful when performing a forensic computer analysis.

Tools Included in Penguin Sleuth Kit

The Penguin Sleuth Kit Base Package contains the following applications:

- **Gentoo Linux 2.6 Kernel** Optimized for Forensics Use
- **XFCE** GUI
- **Apache2** Server
- **MySQL** PHP4
- **Open Office**
- **Gimp** Graphics Program
- **KSnapshot** Screen Capture Program
- **Mozilla**
- **Gnome** CD Master
- **K3b** CD Burner
- **XMMS** Media player
- **Porthole** Gentoo Graphics Package Manager
- **Karchiver** GZIp GUI

It also includes a number of Forensics Tools, including:

- **Sleuth Kit** Forensics Kit
- **Py-Flag** Forensics Browser
- **Autopsy** Forensics Browser for Sleuth Kit
- **dcfldd** DD Imaging Tool command line tool and also works with AIR
- **foremost** Data Carver command line tool
- **Air** Forensics Imaging GUI
- **md5deep** MD5 Hashing Program
- **netcat** Command Line
- **cryptcat** Command Line
- **NTFS** Tools
- **qtparted** GUI Partitioning Tool
- **regviewer** Windows Registry Viewer

Additionally, the CD distribution also contains a number of Security Tools that may aid in incident response and investigation:

- **Etherape** GUI Network Traffic Monitor
- **Clamv** Anti Virus
- **snort** Command Line
- **John the Ripper** Command Line password cracker
- **rkhunter** Command Line
- **Ethereal** Network Traffic Analyzer
- **FWBuilder** GUI Firewall App
- **nessus** network scanner

Forensix

The goal of the Forensix Project (National Science Foundation [NSF] under Grant ANI-0230960—AKA "4N6") is to allow a system to be monitored so that, in the event of a security compromise, it is easy to track the compromise back to its source. The Forensix system requires two machines:

- The potentially insecure "front-line" machine
- A verified secure back-end

Forensix stores information describing system calls in a MySQL database on the back-end and is built on top of SNARE.

Forensix is designed to aid forensic examiners by increasing the efficiency, accuracy, and ease of completing an imaging and analysis assignment. The product aids in ensuring the integrity of image data and introduces a scientific process that creates reproducible analytical results with detailed documentation.

Maresware

Mares and Company (www.maresware.com/) produce a number of programs suited to computer forensics and data analysis. One forensic analyst using Linux is the software suite, Maresware: Linux Forensics.

Maresware: Linux Forensics provides tools for investigating computer records while running under the LINUX operating system on Intel processors. The software aids in the forensic discovery of evidence. Maresware: Computer Forensics software is a suite of tools designed for forensic analysis. The entire package of software applications is available together in Maresware: The Suite.

Major Programs Present in Maresware

The (F) or (D), or both (F,D), after each description designates which package the program is included in: (F)orensic or (D)ata Analysis. The programs unique to this suite, which are not contained in any other packages, include: Boot, Dater, Eventlog, Modify, and Mxcopy. These programs add several functions to your toolkit in the areas of systems maintenance, administrative tasks, and computer security. An asterisk (*) indicates that a similar program exists for LINUX.

- **Bates_no** A unique program for adding identifying numbers to filenames in e-documents. Prepare your records for evidentiary use or simply use as a records management tool.
- **Catalog** Catalogs every file on a Linux file system. Also identifies headers. Similar to Maresware Diskcat (which is DOS based).
- **Hash** Performs MD5 (CRC, or SHA) hash of every file on a drive.
- **Hashcmp** Compares outputs of successive hash runs.
- **Md5** Calculates MD5 hash of a file. MIT's Linux source code for MD5 has been compiled to run on a Linux operating system.
- **Strsrch** Searches files for text strings. Similar to Maresware *Strsrch* (which is DOS based).
- **U_to_A** Converts UNIX text to DOS text.

A complete list of software available from Mares and Company may be found at www.maresware.com/maresware/suite_programs.htm.

Captain Nemo

Captain Nemo is a "Multi Platform File Manager" that allows mapping and access to Novell, NTFS, and Linux drives from any Windows computer. It consists of a file viewer, which enables viewing and access to almost any file type. It also offers a search function, which enables the capability to conduct extensive searches for specific text or files. It will also display the owners and attributes of every file to the limit of the file system that is being analyzed.

The product also allows the forensic investigator to copy files from an imaged Netware/NTFS/Linux partition to the Windows partition (see Table 13.3 for details).

Table 13.3 Supported File Systems with Captain Nemo

File System	Supported File Systems	Remarks
Novell	Netware 3.xx Netware 4.xx Netware 5.xx	File sub allocation is supported. File compression is supported. Long file names are supported. Volumes consisting of several drives aren't supported.
Microsoft Windows	NTFS 2000 XP	File compression is supported. Striping is not supported.
Linux	ext2fs ext3fs	

The Farmer's Boot CD

The Farmer's Boot CD (available for download from www.forensicbootcd.com) leads an emerging trend in Digital Forensics by allowing the previewing of a system before imaging it. This is useful with today's large hard drives because if, during the preliminary review, significant data is discovered then the package switches modes to acquire the image. If on the other hand no data of relevance is discovered, then the forensic analyst can save time and storage space by moving on to the next system.

The primary goal of The Farmer's Boot CD (FBCD) is to provide the forensic analyst with the tools and methodology to securely and promptly sample systems (hard drives, USB drives, digital music devices, digital camera media, and anything else of interest) to aid in the identification of evidence and in making the decision whether or not to proceed to imaging.

As taken from the above Web site, The Farmer's Boot CD allows the forensic analyst to:

- Boot most x86 systems
- Mount file systems in a forensic manner
- Preview data using a single, unified graphical user interface (GUI)
- Acquire media after you've previewed the data and found it pertinent to your investigation
- Automatically mount file systems
- Automatically update superblocks on RAID arrays
- Preview Windows operating systems, viewing important data stores with a simple click of the mouse. This includes viewing the Recycle Bin, reading Windows Event

logs, finding e-mail addresses and URLs in the *pagefile.sys* file, and reviewing web browser (Internet Explorer, Opera, and Mozilla) cache file entries (both cookies and history information).

- Preview Linux operating systems including parsing web browser activity and reviewing important log files, and collect a list of recently used files.
- Use a Windows Explorer-like file manager, point-and-clicking through the directory structure.
- Generate lists of system users and attempt to crack their passwords, for both Windows and Linux operating systems.
- Decrypt EFS-encrypted files found on some Windows operating systems.
- Convert between many different date formats with one click of your mouse: UNIX 32bit, Windows 64bit, human readable, and hexadecimal.
- Catalog a file system, searching for only those files that pertain to your case, by identifying them either by extension or by file header.
- Generate thumbnails for all graphics files on a file system either by extension or by header.
- Preview digital cameras and preview the contents of onboard memory, dumping software information as well as any graphics or video files.
- Authenticate and acquire file systems or devices in simple point-and-click GUIs.
- Generate a catalog of all hardware attached to the system, including serial numbers, make and model information, and settings and configuration information.
- Dump BIOS information of a system, providing you with a file that most practitioners never see. Information obtained is BIOS dependent, but typically includes BIOS version, vendor information, characteristics, universally unique IDs (UUID), and serial numbers.
- Obtain both hard drive and file system information by pointing-and-clicking.
- Generate a list of deleted files for FAT12/16/32, NTFS, and ext2 file system types.
- Capture network traffic and analyze with the provided tools as dictated by your case. This includes HTTP, e-mail, and messenger traffic.
- Quickly preview all sessions found on multi-session CDs.
- Identify data residing in file slack for supported file system types.
- Analyze the Windows registry using the provided tools.

Summary

There are advantages to both Linux- and Windows-based forensic processes. Ultimately, the most successful process will be a result of using each operating system based on its strengths and, where possible, aborting its weaknesses.

Windows live system analysis provides information that is volatile in nature and is lost or destroyed when analyzed off line. The analysis of file slack (the space existing at the end of the file of the last cluster), ADS (NTFS Alternate Data Streams), and system memory can aid in the identification of network logon names, passwords, and other sensitive information. This offers advantages to the forensic investigator since volatile Windows data is a vital source of evidence and leads.

The command *dir /O:d* has an important role during the investigation of a Windows system. Registry is also an important component to be examined during an investigation of a Windows system. Certain keys must be checked and given added importance as they inform the investigator on what services have been running, what services are currently running, and who has logged on to the system over a certain period of time.

Linux imparts greater control, flexibility, and power as a forensics tool than Windows, due to its ability to access devices at a physical layer. The addition of a number of simple utilities make imaging and basic analysis of suspect disks and drives easier, allowing for a simpler imaging process once the forensic analyst becomes familiar with the command-line quirks associated with Linux. To help reduce the issues associated with the Linux command-line, several popular Linux toolkits that provide a GUI interface have been developed.

The addition of NASA's loopback driver (a kernel module) introduces the ability to associate an entire disk image with a single loop device. Alternatively, devices have to be mapped partition by partition. Often raw data with valuable evidence can reside on a disk in an area outside the partition.

The main thing that you need to bring from this chapter is an understanding of the forensic process and how tools can aid you. It is important to understand that it is sometimes necessary and often rewarding to analyze live systems. The important thing to remember here is to minimize the amount of change. Volatile data needs to be collected quickly and you need to document what you have done. This chapter should have provided you with an overview of the tasks that you will need to do in order to conduct a forensic engagement on either Windows or Linux-based systems.

Frequently Asked Questions

Q: If live incident response and forensics always changes the system, how can I do this without corrupting the evidence?

A: This is the primary reason for collecting evidence using a document procedure and process. By maintaining a strong chain of custody over the evidence and documenting each step you take, while trying to minimize any changes you make to the system, you can demonstrate that the additional processes added value without significantly changing the system. Live systems hold valuable information and thus evidence in volatile data. It is not possible to obtain volatile data without making some changes to the system being analyzed, but the benefits far outweigh the risk of damaging the data.

Q: I know what I am doing, why do I have to write it down?

A: Often, court cases and litigation occur months or years after the analysis of the data and collection of the evidence. Keeping good notes mean that the value of the evidence you have collected will be maintained and you are less likely to forget what occurred. Tools such as the Helix Live CD for Windows aid in this process by creating automated logs of the activities that you took within them. The output from these files may then be maintained as part of the evidence.

Q: Examining Windows file slack seems like a lot of effort to get broken bits of files.

A: File fragments are admissible in court. Often, file fragments are not the only thing found in file slack. Windows file slack and RAM slack can contain entire files, e-mail, images, or information from the system memory. Of particular importance, information that is normally never stored on disk at all and is certainly never saved on the system can occasionally be found in file slack. In fact, it is possible to find both user names and passwords within file slack.

Q: I have always used Windows, why should I learn Linux?

A: Linux provides a level of access to the system that is not available on Windows. The secret is to use the tool that provides the best results. It is important to be familiar with both Windows and Linux and there are benefits to both systems.

Q: I have heard that it is always possible to write to a device on Linux. What is the purpose of the read-only option in the mount command then and does it do anything?

A: It is always possible to write to the physical device if you have the right permissions and Linux or UNIX system. An example would be writing to the */dev/hdc* physical drive. When you mount a drive on Linux, the mount point that you set it to can be configured to be read-only. The important thing to remember is not to write to the physical drive and only use the read-only mounted partition. Just because functionality is there does not mean you need to use it. Like everything in a forensic assignment, take due care and record directions.

Chapter 14

Investigating Network Traffic and Investigating Logs

Solutions in this chapter:

- **Overview of the OSI Model**
- **Network Addresses and NAT**
- **Using Log Files to Monitor and Track Activity**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

Networks are exposed to an unending series of attacks and exposures. External threats usually come from the Internet and can be grouped into three broad categories such as denial of service (DoS); using the victim's network as a base of attack, to target other networks; or for threat or alteration of information. Internal threats can come from individuals with legitimate access or from those that are exceeding their level of privilege. This means that a forensic expert must understand how the network works and how to access logged data.

Monitoring for insider and outsider attacks should be a proactive activity, but many attacks may not be detected until after the attack has taken place. This will require an inspection of log files. Only by review of the audit and log files can the attack be analyzed and reconstructed. Many forensic experts don't make full use of log and audit data as they may not understand how to access this information or initially overlook it because of the extensive amount of time it takes to read through thousands upon thousands of log file entries. This chapter will look at how tools like the Log Parser can help with this task and make a tedious task easy.

Overview of the OSI Model

In 1984 the International Standards Organization developed the Open Systems Interconnect (OSI) model. The model is designed to provide order by specifying a specific hierarchy in which each layer builds upon the output of each adjacent layer. The model is still used today as a guide in describing the operation of a networking environment.

Layers of the OSI Model

The seven layers of the OSI model are the physical, data link, network transport, session, presentation, and application layers. Let's take a look at each of these layers.

The Physical Layer

Layer 1 is known as the physical Layer. At Layer 1, bit-level communication takes place. The bits have no defined meaning on the wire, but the physical layer defines how long each bit lasts and how it is transmitted and received. From a forensics standpoint a great deal of sensitive information may be available at the physical layer if no encryption is used.

The Data Link Layer

Layer 2 is known as the data link layer. The data link layer is responsible for formatting and organizing the data before sending it to the physical layer. The data link layer organizes the data into frames. A frame is a logical structure in which data can be placed. When a frame reaches the target device, the data link layer is responsible for stripping off the data frame and

passing the data packet up to the network layer. The data link layer is made up of two sublayers, including the logical link control layer (LLC) and the media access control layer (MAC).

The Network Layer

Layer 3, the network layer, is concerned with logical addressing and routing. The network layer is the home of the Internet Protocol (IP), which makes a best effort at delivery of datagrams from their source to their destination.

The Transport Layer

Layer 4 is known as the transport layer, and ensures completeness by handling end-to-end error recovery and flow control. Transport-layer protocols include TCP, a connection-oriented protocol. It provides reliable communication through the use of handshaking, acknowledgments, error detection, and session teardown and User Datagram Protocol (UDP), a connectionless protocol. It offers speed and low overhead as its primary advantage.

The Session Layer

Layer 5 is known as the session layer. Its functionality is put to use when creating, controlling, or shutting down a TCP session. Items such as the TCP 3-way handshake and the TCP 4-way shutdown occur here. Session-layer protocols include items such as Remote Procedure Call and Structured Query Language.

The Presentation Layer

Layer 6, the presentation layer, is responsible for taking data that has been passed up from lower levels and put it into a format that application layer programs can understand. These common formats include ASCII, EBCDIC, and ANSI.

The Application Layer

Layer 7 is known as the application layer. Recognized as the top layer of the OSI model, this layer serves as the window for application services. The application layer is one that most users are familiar with as it is the home of e-mail programs, FTP, Telnet, web browsers, office productivity suites, and many other applications.

Network Addresses and NAT

In networks that use TCP/IP the IP address scheme is used for logical addressing. A MAC address is considered a physical address, whereas an IP address is considered a logical address. IP addresses are configured as dotted decimal notation. The IPv4 address format is four

decimal numbers that are separated by decimal points. Each of these decimal numbers are one byte in length to allow numbers to range from 0–255.

- **Class A Networks**. Class A networks consist of up to 16,777,214 client devices, and their address range can extend from 1 to 126.
- **Class B Networks**. Class B networks host up to 65,534 client devices and their address range can extend from 128 to 191.
- **Class C Networks**. Class C networks can have a total of 245 devices and their address range can extend from 192 to 223.
- **Class D Networks**. These addresses are reserved for multicasting and can extend from 224 to 239.
- **Class E Networks**. These addresses are reserved for experiential purposes. Their addresses range from 240 to 254

Network Address Translation (NAT) originally was developed because of the explosive growth of the Internet, and with the increase in home and business networks, the number of available IP addresses is simply not enough. Typically, Internet Service Providers (ISPs) allocate only one address to a single customer. Companies can buy lots of addresses, but they must pay for each one. Also direct connectivity to the Internet is not without its dangers. NAT is a cost-effective secure solution. NAT allows a single device, such as a router, to act as an agent between the Internet and the local network. This device or router provides a pool of addresses to be used by your local network.

Network Information-Gathering Tools

Network information-gathering tools are pieces of software that can be used to obtain network data for forensic analysis. These tools generally offer the functionality of a sniffer and an intrusion detection system combined. Sniffers are powerful programs that work by placing the hosts system's network card into promiscuous mode. A network card in promiscuous mode can receive all the data it can see, not just packets addressed to it. Switches segment traffic and know which particular port to send traffic to and block it from all the rest. Although this feature adds much needed performance gains, it does raise a barrier when attempting to sniff all potential switched ports. Forensic analysis will usually require the switch to be configured to mirror a port. Some common network monitoring tools include:

- **NetWitness**. NetWitness is designed to analyze network traffic and monitor it.
- **Netresident Tool**. Considered an advanced network content monitoring program that captures, stores, analyzes, and reconstructs network events such as e-mail messages, Web pages, downloaded files, and other types of network traffic.

- **Infinistream Security Forensics**. A commercial product that builds on sniffer technology by providing high-end tracking of absolutely everything.
- **CA Network Forensics**. Allows the user to uncover and investigate network traffic. This program captures raw network data and uses forensic analysis to check for exploitation, internal data theft, and security violations.
- **Wireshark**. An open source protocol analyzer that can capture traffic in real time.
- **Snort**. A well-known open source IDS product that uses signatures to detect events.

Sniffers

Sniffers operate at the data link layer of the OSI model. This means they do not have to play by the same rules as applications and services that reside further up the stack. Sniffers can capture everything on the wire and record it for later review. They allow the user to analyze all the data that is contained in the packet. A good sniffer product is Wireshark. Not only is it free to use but it works well in both a Windows and Linux environment. You can find more details about the program at the official site, www.wireshark.com. Next, we will discuss the second category of network information gathering tools, intrusion detection.

Intrusion Detection

Intrusion detection systems (IDS) play a second critical role in the protection of the IT infrastructure. Intrusion detection involves monitoring network traffic, detecting attempts to gain unauthorized access to a system or resource, and notifying the appropriate individuals so that counteractions can be taken. Having the ability to analyze intrusions and attacks is a powerful tool. There are four types of logs that would be of interest during a forensic investigation—authentication, application, operating system, and network—but the IDS will be of use primarily with network logs. There are many good IDS systems available. Snort is one example of an IDS with wide industry support.

NOTE

Intrusion detection actually has a rather short history. In 1983 Dr. Dorothy Denning started building on the work of James Anderson and developed the first IDS system for the U.S. government. They were interested in analyzing the audit trails of government mainframe systems to determine misuse.

Snort

Snort is a freeware IDS developed by Martin Roesch and Brian Caswell. It's considered a lightweight network-based IDS that can be set up on a Linux or Windows host. The core program has a command line interface, but there are GUIs that can be used. Snort operates as a network sniffer and logs activity that matches predefined signatures. Signatures can be designed for a wide range of traffic including Internet Protocol, Transmission Control Protocol, User Datagram Protocol, and Internet Control Message Protocol. An intrusion detection system is valuable to a forensic analyst only if the data is reviewed and analyzed.

Unfortunately, sometimes an IDS will produce overwhelming amounts of data that make it difficult to process. To aid in our interpretation of the data, we can use Microsoft Log Parser to take snapshots of our IDS logs and present them in different easy-to-read reports. In this chapter, we will build an example IDS report using nothing more than the power of Log Parser.

Gathering Snort Logs

To process the alert data, we first need a consistent method for gathering the data. Log Parser is an excellent method for managing Snort logs because you can query the file while Snort still has the log open. Many administrators schedule scripts to regularly cycle the Snort logs, but this requires stopping the service to release the file so a script can move it. Using Log Parser, we can use checkpoints to read the most recent data from the file.

Although Snort supports several output formats that Log Parser could use, I have found the CSV format most flexible and consistent. To configure Snort to use the CSV output format, simply add the following line in the snort.conf file:

```
output alert_csv: alert.csv default
```

This configures Snort to create a CSV log file named alert.csv in the configured logs directory using the default output fields. By default the CSV output processor includes these fields:

- timestamp
- sig_generator
- sig_id
- sig_rev
- msg
- proto
- src
- srcport
- dst
- dstport

- ethsrc
- ethdst
- ethlen
- tcpfags
- tcpseq
- tcpack
- tcplen
- tcpwindow
- ttl
- tos
- id
- dgmlen
- iplen
- icmptype
- icmpcode
- icmpid
- icmpseq

Snort CSV logs do not include a header row, so we need a separate file to name each column. To read CSV Snort alerts, you would use a command like this:

```
logparser.exe  le:alert.sql -i:csv -headerRow:off -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss
```

Note that we specify the CSV input format, but instead of using the header row, we specify a header file using the **iHeaderFile** option. We also specify the timestamp format so Log Parser can interpret that field as an actual time stamp rather than a string.

NOTE

All the queries in this chapter use the year in the timestamp date, which Snort does not log by default. To configure Snort to log the year, use the **–y** option when starting Snort. If your timestamps do not include the year, the queries in this chapter will return the error, "Semantic Error: argument of function TO_TIME must be a TIMESTAMP".

Building an Alerts Detail Report

In our IDS report we likely want to view summaries of the alert data such as:

- Most common source IP (Internet Protocol) addresses
- Most common target IP addresses

Using Log Parser's multiplex feature and template output format we can very easily create interactive HTML (Hypertext Markup Language) reports directly from the Snort logs.

Alerts by IP Address

Each IP address in the alerts report shown in Figure 14.1 is a clickable hyperlink that leads to a detail page showing all alerts for that IP address. Using a process similar to that used previously for the alert messages, wecreated a summary page (Figure14.2) and detail pages (Figure 14.3) using a two-pass approach. We repeated this process for both source and destination IP addresses to produce a fully interactive HTML IDS report.

Figure 14.1 Detailed Alert Messages

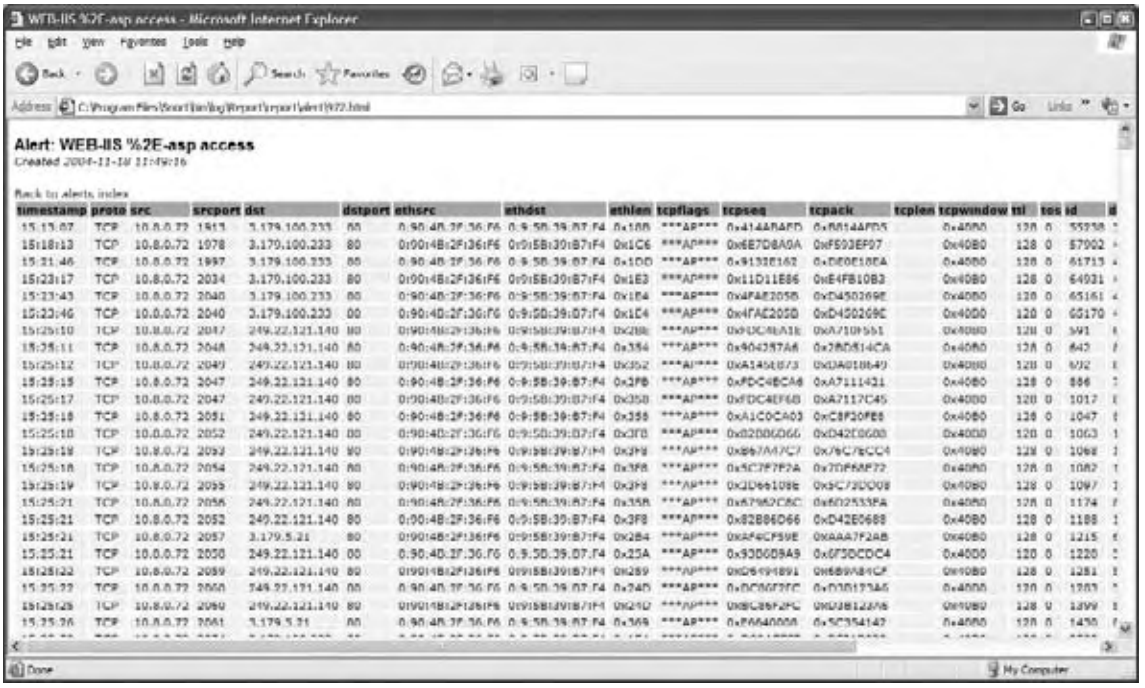

WEB-IIS %2E-asp access - Microsoft Internet Explorer
Alert: WEB-IIS %2E-asp access
timestamp proto src srcport dst dstport ethsrc ethdst ethlen tcpflags tcpseq tcpack tcplen tcpwindow ttl tos id

Figure 14.2 Snort Alerts by Destination IP Address

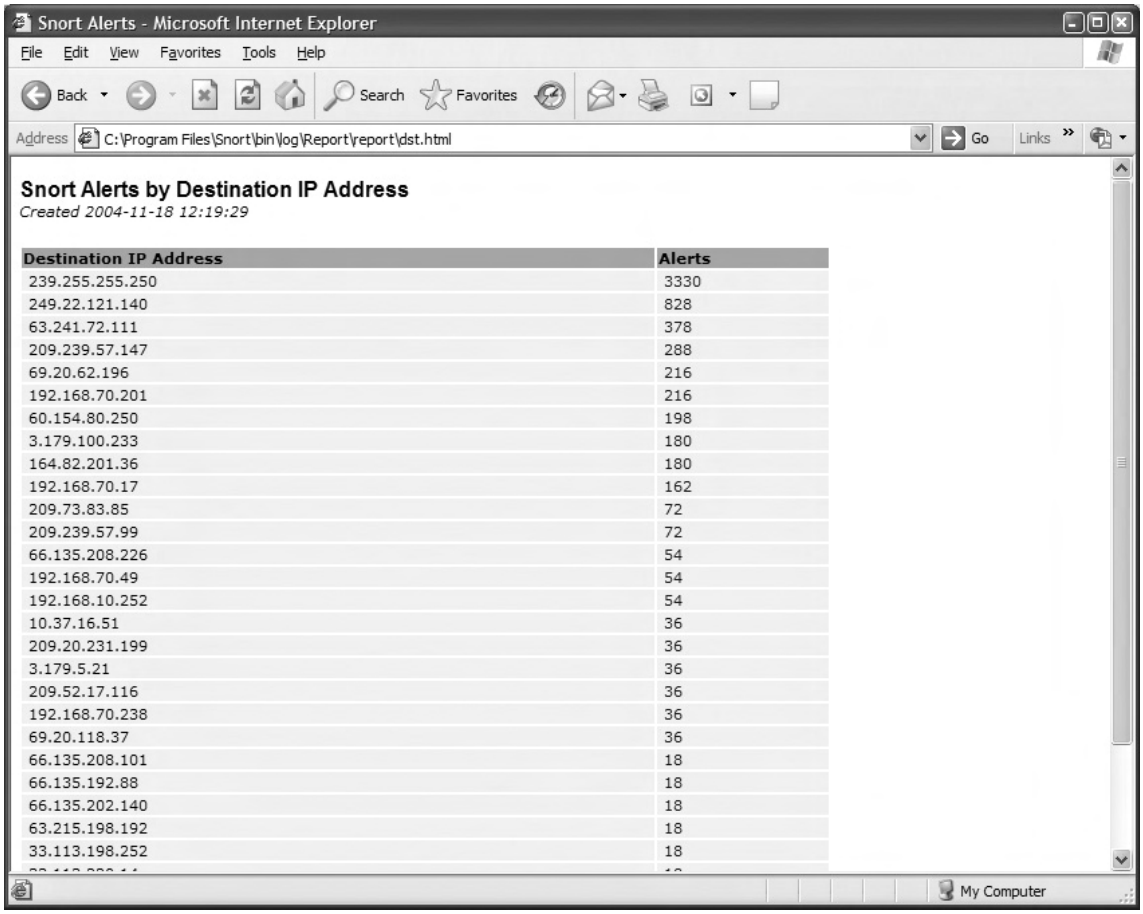
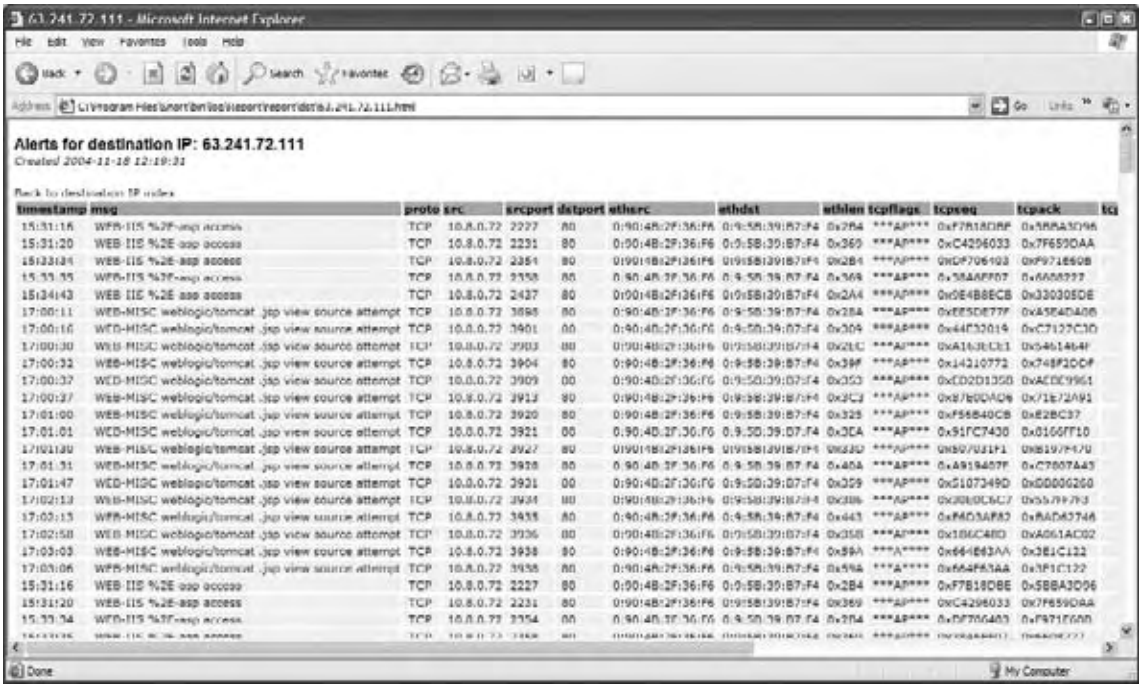


Figure 14.3 IP Address Details


Alerts for destination IP: 63.241.72.111
timestamp msg proto src srcport dstport ethsrc ethdst ethlen tcpflags tcpseq tcpack
15:31:16 WEB-IIS %2F-asp access TCP 10.8.0.72 2227 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2B4 ***AP*** 0xF7B18DBE 0x5BBA3D96
15:31:20 WEB IIS %2E-asp access TCP 10.8.0.72 2231 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x369 ***AP*** 0xC4296033 0x7F659DAA
15:33:34 WEB-IIS %2E-asp access TCP 10.8.0.72 2351 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2B4 ***AP*** 0xDF706103 0xF971E60B
15:33:35 WEB-IIS %2F-asp access TCP 10.8.0.72 2352 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x369 ***AP*** 0x58AAFF07 0x6A58777
15:34:43 WEB IIS %2E asp access TCP 10.8.0.72 2427 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2A4 ***AP*** 0x0E4B8ECB 0x330305DE
17:00:11 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3898 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x284 ***AP*** 0xEE5DE77F 0xA5E4DA06
17:00:16 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3901 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x304 ***AP*** 0x44F53014 0xC7125C3D
17:00:30 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3903 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2EC ***AP*** 0xA163E3E1 0x5461464F
17:00:32 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3904 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x39F ***AP*** 0x14210772 0x748F2DDF
17:00:37 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3909 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x353 ***AP*** 0xCD2D135D 0xACDE9961
17:00:37 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3913 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2C2 ***AP*** 0x87B0DAD6 0x71E72A91
17:01:00 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3920 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x325 ***AP*** 0xF56B40CB 0xE2BC37
17:01:01 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3921 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x3DA ***AP*** 0x91FC7430 0x0166FF10
17:01:30 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3927 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x33D ***AP*** 0xB07U31F1 0xB19/F47U
17:01:31 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3928 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x40A ***AP*** 0xA919407F 0xC7807A43
17:01:47 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3931 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x359 ***AP*** 0x5107349D 0xD0006260
17:02:13 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3934 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x3B6 ***AP*** 0x3B0C5C17 0x55/97F3
17:03:13 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3935 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x443 ***AP*** 0xF4D3AF83 0xBAD43746
17:02:58 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3936 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x358 ***AP*** 0x186C4B0 0xA061AC02
17:03:03 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3938 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x39A ***AP*** 0x664B63AA 0x3E1C122
17:03:06 WEB-MISC weblogic/tomcat .jsp view source attempt TCP 10.8.0.72 3939 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x39A ***AP*** 0x664F63AA 0x3E1C122
15:31:16 WEB-IIS %2E asp access TCP 10.8.0.72 2227 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2B4 ***AP*** 0xF7B18DBE 0x5BBA3D96
15:31:20 WEB-IIS %2E-asp access TCP 10.8.0.72 2231 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x369 ***AP*** 0xC4296033 0x7F659DAA
15:33:34 WEB-IIS %2F-asp access TCP 10.8.0.72 2354 80 0:90:4B:2F:36:F6 0:9:5B:39:B7:F4 0x2B4 ***AP*** 0xDF706403 0xF971E60B
Done
My Computer

At this point, you can run the entire report with these Log Parser commands:

```
logparser.exe  le:Ch14Alerts-Indexsql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14Alerts-Index.tpl

logparser.exe  le:Ch14Alerts-DetailHeader.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14Alerts-DetailHeader.tpl

logparser.exe  le:Ch14Alerts-Detail.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14Alerts-Detail.tpl
- leMode:0

logparser.exe  le:Ch14SrcIP-Index.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14SrcIP-Index.tpl

logparser.exe  le:Ch14SrcIP-DetailHeader.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14SrcIP-DetailHeader.tpl

logparser.exe  le:Ch14SrcIP-Detail.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14SrcIP-Detail.tpl
- leMode:0

logparser.exe  le:Ch14DstIP-Index.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14DstIP-Index.tpl
```

```
logparser.exe  le:Ch14DstIP-DetailHeader.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14DstIP-DetailHeader.tpl
logparser.exe  le:Ch14DstIP-Detail.sql -i:csv -iHeaderFile:AlertHeader.csv
-iTsFormat:mm/dd/yy-hh:mm:ss -headerRow:off -o:tpl -tpl:Ch14DstIP-Detail.tpl
- leMode:0
```

Building an Alerts Overview Report

Now that we have a detailed alerts report, we might want to build a summary index page. This page should include links to the detailed reports and also display graphs and short summaries of the data to get a quick overview of the network

Since the summary report includes only one table of data with multiple graphs, I placed the entire HTML markup in a single template file, Ch14Summary-Index.tpl. This template creates the HTML for the report that includes references to the individual graphs. The query for the top alerts is similar to that used for the alerts index page, but includes only the top 10 records:

```
---Ch14Summary-Index.sql---
SELECT TOP 10
  sig_id,
  msg,
  Count(msg) as Alerts
INTO report\index.html
FROM alert.csv
GROUP BY msg, sig_id
ORDER BY Alerts DESC
---Ch14Summary-Index.sql---
```

The query for the pie graph is similar, but does not include the actual message, and this time processes all records:

```
---Ch14Summary-GraphTopAlerts.sql---
SELECT
  sig_id,
  Count(msg) as Alerts
INTO report\AlertsTopAlerts.gif
FROM alert.csv
GROUP BY sig_id
ORDER BY Alerts DESC
---Ch14Summary-GraphTopAlerts.sql---
```

Finally, there are three queries for the remaining graphs:

```
---Ch14Summary-GraphTopSrcIPs.sql---
SELECT
  src,
  Count(msg) as Alerts
INTO report\AlertsTopSrcIPs.gif
FROM alert.csv
GROUP BY src
ORDER BY Alerts DESC
---Ch14Summary-GraphTopSrcIPs.sql---
---Ch14Summary-GraphAlertsPerHour.sql---
SELECT
  Count(*) as Alerts
USING QUANTIZE(timestamp,360) as Hour
INTO report\AlertsByHour.gif
FROM alert.csv
GROUP BY Hour
---Ch14Summary-GraphAlertsPerHour.sql---
---Ch14Summary-GraphTopDstPorts.sql---
SELECT TOP 5
  STRCAT(STRCAT(TO_STRING(dstport),' - '), proto) AS Destination,
  Count(*) as Alerts
  USING dst as DestinationPort
INTO report\AlertsTopDstPorts.gif
FROM alert.csv
GROUP BY Destination
ORDER BY Alerts DESC
---Ch14Summary-GraphTopDstPorts.sql---
```

Finally, we can generate the entire index page with these commands:

```
logparser.exe  le:Ch14Summary-Index.sql -i:csv
-iHeaderFile:AlertHeader.csv -iTsFormat:mm/dd/yy-hh:mm:ss
-headerRow:off -o:tpl -tpl:Ch14Summary-Index.tpl
logparser.exe  le:Ch14Summary-GraphTopAlerts.sql -i:csv
-iHeaderFile:AlertHeader.csv -iTsFormat:mm/dd/yy-hh:mm:ss
-headerRow:off -o:chart -chartType:Pie3D -groupSize:350×190
-values:OFF -chartTitle:"" -categories:OFF
logparser.exe  le:Ch14Summary-GraphTopSrcIPs.sql -i:csv
-iHeaderFile:AlertHeader.csv -iTsFormat:mm/dd/yy-hh:mm:ss
-headerRow:off -o:chart -chartType:Pie -groupSize:300×150
-values:OFF -chartTitle:"" -categories:OFF
```

```
logparser.exe  le:Ch14Summary-GraphAlertsPerHour.sql -i:csv
-iHeaderFile:AlertHeader.csv -iTsFormat:mm/dd/yy-hh:mm:ss
-headerRow:off -o:chart -chartType:smoothline -groupSize:300×150
-values:OFF -chartTitle:“” -categories:OFF
logparser.exe  le:Ch14Summary-GraphTopDstPorts.sql -i:csv
-iHeaderFile:AlertHeader.csv -iTsFormat:mm/dd/yy-hh:mm:ss
-headerRow:off -o:chart -chartType:BarStacked -groupSize:300×150
-values:OFF -chartTitle:“”
```

The final result is a fully interactive IDS report using nothing more than Log Parser.

Monitoring User Activity

Monitoring user activity is an important part of the forensic process. Unusual user activity may be just an indicator of system problems, or it may be a real security issue. By analyzing the user activity reported in your system logs, you can determine authentication problems and hacking activity.

NOTE

When working with older systems it is important to remember that security Auditing is not enabled on Windows 2000 systems by default.

An important aspect in determining relevant events in your system logs is determining which behavior is benign and which behavior indicates trouble. For example, if a user fails to authenticate one time in a week span, you can be pretty certain that the user simply mistyped his or her password and is not trying to hack into your system. Alternatively, if a user has two failed authentication attempts hourly for an extended period you should look more deeply into the problem. We will be covering these types of events and going over how to make the critical determination as to whether a behavior is benign or a warning sign. Security is, and always should be, a top priority for anyone working on key infrastructure systems. Tools like Log Parser can help you identify events that could indicate a potential hack attempt or even just a user attempting to look at something he or she shouldn't be looking at. By enabling file access auditing and using Log Parser to examine the results, you can even track down potential security issues at the file level. Digging through these same log files to track down this data manually is both tedious and unnecessary with the advanced capabilities of Microsoft's Log Parser utility.

Tracking Authentication Failures

When performing regular security audits of your servers, it is always important to know how many failed authentication attempts have occurred. This helps you in several ways. First, you are able to determine specific users who may have forgotten their password and failed to request a new one. Second, and more importantly, you may be able to detect potential hacking attempts against your server by analyzing the results and noting an increase in failed logon attempts either for a single user or for multiple users. Finally, you may be able to determine problems with the system that are causing authentication failures by multiple users at a specific time. These problems could relate to network connectivity, system utilization, or a failure of an authentication server.

Listing Failed Logons

Let's return again to the Log Parser to see how easily you can use your event logs to detail out the failed logons over a specific timeframe. The following command and query can be used to list all failed logons on a specific date:

```
logparser.exe  le:Ch14ListingFailedLogons.sql -i:EVT -o:datagrid
---Ch14ListingFailedLogons.sql---
SELECT
  timegenerated AS LogonTime,
  extract_token(strings, 0, '|') AS UserName
FROM Security
WHERE EventID IN (529;
  530;
  531;
  532;
  533;
  534;
  535;
  537;
  539)
  AND to_string(timegenerated,'yyyy-MM-dd HH:mm:ss') like '2004-09%'
---Ch14ListingFailedLogons.sql---
```

By running this query, you are able to see how many authentication failures have occurred on the date you specified, as well as which user IDs experienced failures.

In this query, we begin by specifying that we want to obtain the timestamp for the event from the *timegenerated* field. From there, the query gets a little trickier than usual. For failed logon events, the UserID for the event itself is always SYSTEM. In order to get the actual ID that the logon was attempted with, we need to parse out some data from the description

or *Strings* field. To do so, we use the EXTRACT_TOKEN function to tokenize the string using the pipe (|) character and specify that we want the first or 0 token, which is the UserName.

Next we specify that we wish to pull this data from the current Security log. Our WHERE clause specifies a specific range of security events that indicate logon failures, and we use this to limit our data to those specific events. A description of each of these events is detailed in Table 14.1. In the WHERE clause, we also specify the date stamp that we're looking for in the same manner as the previous query. We also need to specify the input and output format, so we add this to the end of the command line using the syntax **-i:EVT –o: datagrid**.

Table 14.1 Failed Logon EventIDs

EventID	Description
529	The logon attempt was made with an unknown username or a known username with a bad password.
530	The user account tried to log on outside the allowed time.
531	A logon attempt was made by using a disabled account.
532	A logon attempt was made by using an expired account.
533	The user is not allowed to log on at this computer.
534	The user attempted to log on with a logon type that is not allowed, such as network, interactive, batch, service, or remote interactive.
535	The password for the specified account has expired.
537	The logon attempt failed for other reasons.
539	The account was locked out at the time the logon attempt was made. This event is logged when a user or computer attempts to authenticate with an account that has been previously locked out.

Identifying Single versus Multiple Failed Logons

When working with failed logons, it sometimes helps to be able to quickly determine how prevalent logon failures are over a specific time period. In the preceding section, we went over how to list failed logons, but using this query requires manual determination of how frequently a specific user ID is experiencing problems.

You can use Log Parser to automatically tally up the number of failed logons for each user by utilizing a slightly different query. The following command and query is one example of how this can be accomplished:

```
logparser.exe  le:Ch14SingleVsMltplFailedLogons.sql -i:EVT -o:datagrid
---Ch14SingleVsMltplFailedLogons.sql---
SELECT
  extract_token(strings, 0, '|') AS UserName,
  count(*) AS Number_of_Events
FROM Security
WHERE EventID IN (529;
  530;
  531;
  532;
  533;
  534;
  535;
  537;
  539)
  AND to_string(timegenerated,'yyyy-MM-dd HH:mm:ss') like '2004-09%'
GROUP BY UserName
HAVING Number_of_Events>2
---Ch14SingleVsMltplFailedLogons.sql---
```

In this query, we are doing a bit more work up front, but minimizing the manual effort required to identify multiple login failures. We start out with the EXTRACT_TOKEN function parsing out the actual userid with the failed logon. Next, we perform a count of the number of events and display that value as Number_of_Events. We then identify the Security event log as our data source and add in a WHERE clause specifying the events that we want to watch for, as well as the specific date that we want the data from.

Finally, we add a GROUP BY statement to organize our results and also add in a HAVING statement. To limit the data to show us data only from events that occur multiple times, we specify Number_of_Events>2 in the HAVING statement.

Identifying Brute Force Attacks

When dealing with system security, a major concern is the possibility of a *brute force* attack on your system. The determining factor as to whether or not an attack is a brute force attack is the manner in which the attack occurs. When attempt after attempt after attempt is made to perform a specific attack action, the overall event is considered a brute force attack because the attacker is trying to break into the system using brute force.

Log Parser can help you quickly analyze your security event logs and determine whether or not one of these brute force attacks is potentially occurring. By creating queries that watch for the behavior of a brute force attack, you may be able to stop an attack before it is successful.

Identifying a Brute Force Authentication Attack

A great example of a brute force attack is one in which someone is trying to guess the password for one of your users. Let's say that the attacker knows from previous experience that you have a policy in place that disables an account if three unsuccessful logon attempts are made within a one-hour window by the same UserID. With this in mind, the attacker is not going to risk locking out the account that he or she is trying to hack. However, the attacker will want to maximize his or her available time and attempt the maximum number of attacks that he or she can within a specific time period.

A quick way to check for this type of activity is to use the same query listed earlier for identifying multiple failed logons. With a little modification, this query can perform an intelligent scan for brute force logon attempt behavior. The original query is as follows:

```
logparser.exe  le:Ch14SingleVsMltplFailedLogons.sql -i:EVT -o:datagrid
---Ch14SingleVsMltplFailedLogons.sql---
SELECT
  extract_token(strings, 0, '|') AS UserName,
  count(*) AS Number_of_Events
FROM Security
WHERE EventID IN (529;
  530;
  531;
  532;
  533;
  534;
  535;
  537;
  539)
  AND to_string(timegenerated,'yyyy-MM-dd HH:mm:ss') like '2004-09%'
GROUP BY UserName
HAVING Number_of_Events>2
---Ch14SingleVsMltplFailedLogons.sql---
```

As is, this query works great for determining multiple failed logon events on the same day. However, that behavior in and of itself does not signify that a brute force attack is occurring. In order to make this query more effective, we need to add the behavior signature of a brute force attack and get a little more detail.

The best way to add that behavior search into our query is to go back to what the attacker has determined your security policy is. The account will be locked out if more than two logon attempts fail within a one-hour time period by the same UserID. So, to maximize his or her potential attack time, the attacker will probably try to perform two logon attempts,

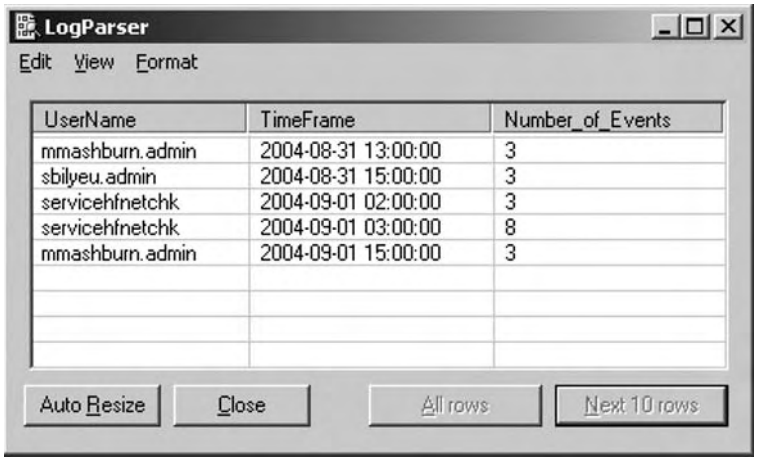
determine whether or not they fail, then wait an hour and try twice more. So what we would look for is three or more failed logon events throughout the day and obtain the timeframe in which those events occurred.

To query based on this information, we need to modify our original query as follows:

```
logparser.exe  le:Ch14BruteForceAttack.sql -i:EVT -o:datagrid
---Ch14BruteForceAttack.sql---
SELECT
  extract_token(strings, 0, '|') AS UserName,
  quantize(timegenerated, 3600) AS TimeFrame,
  count(*) AS Number_of_Events
FROM new_security.evt
WHERE EventID IN (529;
  530;
  531;
  532;
  533;
  534;
  535;
  537;
  539)
GROUP BY UserName,
  timeframe
HAVING Number_of_Events>2
---Ch14BruteForceAttack.sql---
```

The only real changes in this query were the removal of the date search and the addition of a TimeFrame value based on the quantized timegenerated field. When running this query, we will be presented with a datagrid showing the number of failure events for each username within a one-hour timeframe. A quick glance at the output shown in Figure 14.4 shows you the accounts that could be experiencing a brute force attack. Following up with the user to find out more detail on what is going on would be a worthwhile practice when this type of activity is found.

Figure 14.4 Brute Force Activity



Tracking Security Policy Violations

Most companies have developed some form of security policy that addresses concerns around account and password security, permissible system access, unacceptable behavior, and such. These policies dictate the specific rules that each user must follow when accessing information technology (IT) systems.

A useful activity that admins typically perform is a regular audit to ensure that corporate security policies are being adhered to. Some of the information useful for an audit of this nature can be found in the Windows event logs. A great example is tracking user logons for desktop systems. By correlating all your event log data from multiple systems and digging through them with Log Parser, you can determine whether or not users are violating a policy restricting them to log on to only one machine at a time.

Determining Logon/Logoff Behavior

Another example that requires no log correlation is determining whether users typically are logging off after their work day or leaving their systems logged in. Some companies do have this behavior documented as a violation of their corporate security policy.

To gather this data with Log Parser, first we'll need to determine all the logoff events, then check to see if there is a logon event for the same day that does not have an associated logoff event for the same SID. The following command and query can help us determine this behavior:

```
logparser.exe  le:Ch14LogonLogoffBehavior.sql -i:EVT -o:datagrid
---Ch14LogonLogoffBehavior.sql---
SELECT
```

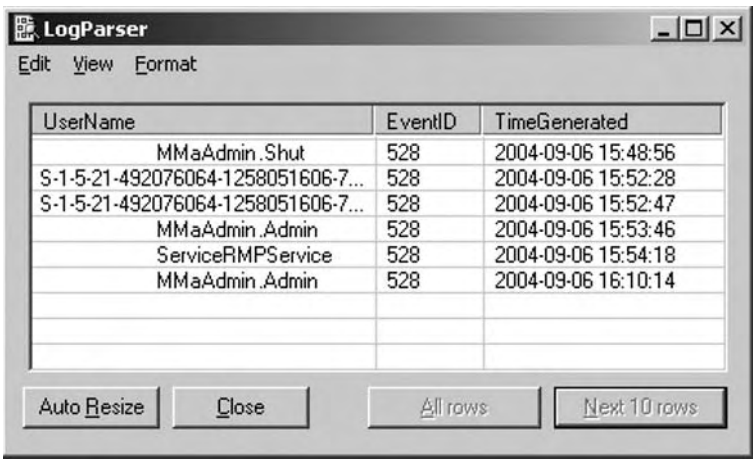
```
  resolve_sid(sid) AS UserName,
  eventid,
  timegenerated
FROM Security
WHERE eventid='528'
  AND to_date(timegenerated)='2004-09-06 00:00:00'
  AND sid not in
  (
  SELECT DISTINCT
    sid
  FROM Security
  WHERE eventid='538'
    AND to_date(timegenerated)='2004-09-06 00:00:00'
  )
---Ch14LogonLogoffBehavior.sql---
```

In this query, we need to work backward a little bit to get the information we need. First, we determine which data we want in our report. In this case, we should gather the userID, the eventID that we're working with, and the time the event occurred. Next, we'll specify the security event log as our data source.

Our creation of the WHERE clause is the more difficult part of this query. The first two items are fairly simple; the eventID has to be 528 for a successful logon and the date that we want information about has to be specified. Also, we'll need to do a *select within a select* to limit our incoming data. Specifically, we want data where the SID with the 528 event does not exist in a query of the 538 events for the same date. You should see results similar to those shown in Figure 14.5.

By performing our query in this manner, we are able to isolate those users who have logged on at some point during a specific day, but do not have an associated logoff event for

Figure 14.5 Logon Events with No Logoff Event



the same day. This will not catch all users who log on without logging off, as it does not take into account the possibility of multiple logon/logoff events for each user in a single day. However, it can help minimize the number of security policy violations of this type.

Auditing Successful and Unsuccessful File Access Attempts

Another important dataset to examine when looking at overall system security is the audit log for successful and unsuccessful file access attempts. Using NTFS (NT file system), you can enable auditing at the file system level for individual files or directories. This allows you, as an administrator, to determine whether or not someone has accessed or attempted to access critical data.

Often, file access auditing is disabled because so much data is generated out of the audit logs that it is almost impossible to sort through the information available and extract events that are important to be aware of. This is where Log Parser can help in your efforts to ensure that your systems are well secured.

Auditing Unsuccessful File Access Attempts

When auditing is enabled for NTFS objects, Windows adds events to the Security event log to indicate the objects that are accessed. As part of this event entry, information is recorded as to the object that was accessed, the user accessing the object, and the date/time that the object was accessed. You can use Log Parser to quickly collect the events around these object access attempts.

Obviously the area more administrators are interested in is the objects that a user tries to access when they're not supposed to. With a good security implementation, this typically results in an unsuccessful file access attempt. Using Log Parser to scan through the security logs, you can quickly identify these events and export them for analysis. The command and query shown here will identify all of the unsuccessful file access events in your security log:

```
logparser.exe  le:Ch14UnsuccessfulFileAccess.sql -i:EVT -o:datagrid
---Ch14UnsuccessfulFileAccess.sql---
SELECT
  timegenerated AS EventTime,
  extract_token(strings, 8, '|') AS UserName,
  extract_token(strings, 2, '|') AS File
FROM Security
WHERE EventID = '560'
  AND EventTypeName = 'Failure Audit event'
  AND extract_token(strings, 1, '|') like 'File'
---Ch14UnsuccessfulFileAccess.sql---
```

When this query runs, it looks for any events in the security log with an eventID of 560 (Success Audit), but further refines that data to events with a type of Failure Audit event. It also

parses through the associated event description to ensure that we look only at files, not directories. From that subset, Log Parser extracts the time of the event, the username associated with the event, and the filename. All this data is then displayed in a datagrid format. You can take this further by having Log Parser export the data into an XML file or some other format for evidence gathering needs.

Auditing Successful File Access Attempts

Many administrators often fail to monitor successful file access attempts in addition to the unsuccessful attempts. Although the most important events to watch for are indeed failures, it is also important not to lose sight of the auditing capabilities available for successful file access attempts.

A scenario where this can be used will help. For example, let's say that you are an administrator for a large corporation with a legal department. The department currently is engaged in putting together some information around a very important case and is storing this information in the department's shared drive. In the course of events, word reaches the legal department that somehow the legal department for the other side of the case has gotten hold of the strategy that your company's legal department was going to use and is putting together a rebuttal. The only people who should have had access to the files containing this information are the members of the legal department. They ask you to track down who has accessed those files.

If you have planned ahead for this eventuality, you have already enabled successful file access auditing for the legal department's shared drive. Using Log Parser, you can easily extract the information that the legal department needs and provide it to them quickly. The following script illustrates how this is done:

```
logparser.exe  le:Ch14SuccessfulFileAccess.sql -i:EVT -o:datagrid
---Ch14SuccessfulFileAccess.sql---
SELECT
  timegenerated AS EventTime,
  extract_token(strings, 8, '|') AS UserName,
  extract_token(strings, 2, '|') AS File
FROM Security
WHERE EventID = '560'
  AND EventTypeName = 'Success Audit event'
  AND extract_token(strings, 1, '|') like 'File'
---Ch14SuccessfulFileAccess.sql---
```

If you look closely at this code, you'll see that this is almost the same as the previous script you saw on auditing unsuccessful file access attempts. The only real difference is the change in the event type name. Instead of "Failure Audit event" we're now looking for "Success Audit event." Running this script will extract all the successful file access events from your security log. You could further refine this to search for events that occur with a specific directory path or filename to make the amount of data to look through smaller.

Summary

In this chapter, we have gone over many of the features available in Log Parser for working with Windows event logs. Through the sections we've covered, we have looked into the Security, System, and Application event logs and used many sample Log Parser queries to pull data from these logs for analysis.

There are a few key points you should remember when working with these logs:

- Differentiating between single and multiple events is a critical feature in Log Parser for working with the Security event log.
- Differentiating between benign events and actual errors or warnings is very important for having an accurate view of your system environment.
- When working with applications and the Application event logs, it is critical that you have a thorough understanding of the specific application you are working with.

Keeping these points in mind will help you make the most of the Log Parser utility. You'll find that as you practice using this tool, it will be a very important part of your overall systems administrator toolbox. Log Parser can help you a great deal in conserving time and effort when you need to parse through large amounts of Windows event log data.

Frequently Asked Questions

Q: Why was the OSI model designed?

A: The OSI model was designed to bring a structured order to the world of networking.

Q: How can I easily remember the layers of the OSI model?

A: One way to remember the layers in their proper order is by **P**lease **D**on't **T**hrow **S**ausage **P**izza **A**way.

Q: Are all layers of the OSI model designed to perform one specific function?

A: No, actually the data-link layer is composed of two sublayers that include the logical link control (LLC) and the media access control (MAC) sublayers.

Q: What is the purpose of NAT?

A: To provide multiple hosts the use of one shared public IP address.

Q: Does NAT have any role in security?

A: Yes, as NAT performs a translation between public and private addresses the internal address scheme is hidden from the public Internet.

Q: What is one of the most widely used information gathering tools?

A: A sniffer. Tools like Wireshark can be used to get a good view of what types of traffic are on the network.

Q: Why monitor user activity?

A: Because we must try to make the distinction between repetitive mistakes and malicious attacks.

Q: Should user activity be examined just after something has gone wrong?

A: No, the best defense is a proactive defense. This means regular periodic monitoring.

Chapter 15

Router Forensics and Network Forensics

Solutions in this chapter:

- Network Forensics
- Searching for Evidence
- An Overview of Routers
- Hacking Routers
- Router Attacks
- Investigation of Routers
- Incident Forensics

- ☑ Summary
- ☑ Frequently Asked Questions

Introduction

This chapter examines router and network forensics. This chapter is important as many attacks will require the analyst to look for information in the router or require network forensics. This requires you to have an understanding of routers and their architecture. It is important to understand where they reside within the OSI model and what role they play within network communications.

Anytime you work with forensic evidence it is critical that the concept of chain of custody be understood. How evidence is handled, stored, accessed, and transported is critical, because if basic control measures are not observed the evidence may be ruled inadmissible in court.

Network Forensics

Network forensics can best be defined as the sniffing, recording, and analysis of network traffic and events. Network forensics are performed in order to discover the source of security incidents and attacks or other potential problems. One key role of the forensic expert is to differentiate repetitive problems from malicious attacks.

The Hacking Process

The hacking process follows a fixed methodology. The steps a hacker follows can be broadly divided into six phases:

1. Reconnaissance
2. Scanning and enumeration
3. Gaining access
4. Escalation of privilege
5. Maintaining access
6. Covering tracks and placing backdoors

The Intrusion Process

Reconnaissance is considered the first preattack phase. The hacker seeks to find out as much information as possible about the victim. The second preattack phase is scanning and enumeration. At this step in the methodology, the hacker is moving from passive information gathering to active information gathering. Access can be gained in many different ways. A hacker may exploit a router vulnerability or maybe social engineer the help desk into giving him a phone number for a modem. Access could be gained by finding vulnerability in the web server's software. Just having the access of an average user account probably won't give the attacker very much control or access to the network. Therefore, the attacker will attempt

to escalate himself to administrator or root privilege. Once escalation of privilege is complete the attacker will work on ways to maintain access to the systems he or she has attacked and compromised. Hackers are much like other criminals in that they would like to make sure and remove all evidence of their activities, which might include using root kits to cover their tracks. This is the moment at which most forensic activities begin.

Searching for Evidence

You must be knowledgeable of each of the steps of the hacking process and understand the activities and motives of the hacker. You many times will be tasked with using only pieces of information and playing the role of a detective in trying to reassemble the pieces of the puzzle. Information stored within a computer can exist in only one or more predefined areas. Information can be stored as a normal file, deleted file, hidden file, or in the slack or free space. Understanding these areas, how they work, and how they can be manipulated will increase the probability that you will find or discover hidden data. Not all suspects you encounter will be super cyber criminals. Many individuals will not hide files at all; others will attempt simple file hiding techniques. You may discover cases where suspects were overcome with regret, fear, or remorse, and attempted to delete or erase incriminating evidence after the incident. Most average computer users don't understand that to drop an item in the recycle bin doesn't mean that it is permanently destroyed.

One common hiding technique is to place the information in an obscure location such as C:\winnt\system32\os2\dll. Again, this will usually block the average user from finding the file. The technique is simply that of placing the information in an area of the drive where you would not commonly look. A system search will quickly defeat this futile attempt at data hiding. Just search for specific types of files such as bmp, tif, doc, and xls. Using the search function built into Windows will help quickly find this type of information. If you are examining a Linux computer, use the grep command to search the drive.

Another technique is using file attributes to hide the files or folders. On a Macintosh computer, you can hide a file with the ResEdit utility. In the wonderful world of Windows, file attributes can be configured to hide files at the command line with the attrib command. This command is built into the Windows OS. It allows a user to change the properties of a file. Someone could hide a file by issuing attrib +h secret.txt. This command would render the file invisible in the command line environment. This can also be accomplished through the GUI by right-clicking on a file and choosing the hidden type.

Would the file then be invisible in the GUI? Well, that depends on the view settings that have been configured. Open a browse window and choose tools/folder options/view/show hidden files; then, make sure Show Hidden Files is selected. This will display all files and folders, even those with the +h attribute set. Another way to get a complete listing of all hidden files is to issue the command attrib /s > attributes.txt from the root directory.

The attrib command lists file attributes, the /s function list all files in all the subdirectories, and > redirects the output to a text file. This text file can then be parsed and placed in a spreadsheet for further analysis. Crude attempts such as these can be quickly surmounted.

An Overview of Routers

Routers are a key piece of networking gear. Let's know the role and function of a router.

What Is a Router?

Routers can be hardware or software devices that route data from a local area network to a different network. Routers are responsible for making decisions about which of several paths network (or Internet) traffic will follow. If more than one path is available to transmit data, the router is responsible for determining which path is the best path to route the information.

The Function of a Router

Routers also act as protocol translators and bind dissimilar networks. Routers limit physical broadcast traffic as they operate at layer 3 of the OSI model. Routers typically use either link state or hop count based routing protocols to determine the best path.

The Role of a Router

Routers are found at layer three of the OSI model. This is known as the networking layer. The network layer provides routing between networks and defines logical addressing, error handling, congestion control, and packet sequencing. This layer is concerned primarily with how to get packets from network A to network B. This is where IP addresses are defined. These addresses give each device on the network a unique (logical) address. Routers organize these addresses into classes, which are used to determine how to move packets from one network to another. All types of protocols rely on routing to move information from one point to another. This includes IP, Novell's IPX, and Apple's DDP. Routing on the Internet typically is performed dynamically; however, setting up static routes is a form of basic routing. Dynamic routing protocols constantly look for the best route to move information from the source to target network.

Routing Tables

Routers are one of the basic building blocks of networks, as they connect networks together. Routers reside at layer 3 of the OSI model. Each router has two or more interfaces. These interfaces join separate networks together. When a router receives a packet, it examines the IP address and determines to which interface the packet should

be forwarded. On a small or uncomplicated network, an administrator may have defined a fixed route that all traffic will follow. More complicated networks typically route packets by observing some form of metric. Routing tables include the following type of information:

- **Bandwidth** This is a common metric based on the capacity of a link. If all other metrics were equal, the router would choose the path with the highest bandwidth.
- **Cost** The organization may have a dedicated T1 and an ISDN line. If the ISDN line has a higher cost, traffic will be routed through the T1.
- **Delay** This is another common metric, as it can build on many factors including router queues, bandwidth, and congestion.
- **Distance** This metric is calculated in hops; that is, how many routers away is the destination.
- **Load** This metric is a measurement of the load that is being placed on a particular router. It can be calculated by examining the processing time or CPU utilization.
- **Reliability** This metric examines arbitrary reliability ratings. Network administrators can assign these numeric values to various links.

By applying this metric and consulting the routing table, the routing protocol can make a best path determination. At this point, the packet is forwarded to the next hop as it continues its journey toward the destination.

Router Architecture

Router architecture is designed so that routers are equipped to perform two main functions: process routable protocols and use routing protocols to determine best path. Let's start by reviewing routable protocols. The best example of a routed protocol is IP. A very basic definition of IP is that it acts as the postman of the Internet—its job is to organize data into a packet, which is then addressed for delivery. IP must place a target and source address on the packet. This is similar to addressing a package before delivering it to the post office. In the world of IP, the postage is a TTL (Time-to-Live), which keeps packets from traversing the network forever. If the recipient cannot be found, the packet can eventually be discarded.

All the computers on the Internet have an IP address. If we revert to our analogy of the postal system, an IP address can be thought of as the combination of a zip code and street address. The first half of the IP address is used to identify the proper network; the second portion of the IP address identifies the host. Combined, this allows us to communicate with any network and any host in the world that is connected to the Internet. Now let us turn our attention to routing protocols.

Routing Protocols

Routing protocols fall into two basic categories, static and dynamic. Static, or fixed, routing is simply a table that has been developed by a network administrator mapping one network to another. Static routing works best when a network is small and the traffic is predictable. The big problem with static routing is that it cannot react to network changes. As the network grows, management of these tables can become difficult. Although this makes static routing unsuitable for use on the Internet or large networks, it can be used in special circumstances where normal routing protocols do not function well.

Dynamic routing uses metrics to determine what path a router should use to send a packet toward its destination. Dynamic routing protocols include Routing Information Protocol (RIP), Border Gateway Protocol (BGP), Interior Gateway Routing Protocol (IGRP), and Open Shortest Path First (OSPF). Dynamic routing can be divided into two broad categories: link-state or distance vector dynamic routing protocols, which are discussed in greater detail later in the chapter.

RIP

RIP is the most common routing protocol that uses a hop count as its primary routing metric. RIP is considered a distance vector protocol. The basic methodology of a distance vector protocol is to make a decision on what is the best route by determining the shortest path. The shortest path is commonly calculated by hops. Distance vector routing is also called routing by rumor.

OSPF

OSPF is the most common link state routing protocol and many times, it is used as a replacement to RIP. Link state protocols are properly called Dijkstra algorithms, as this is the computational basis of their design. Link state protocols use the Dijkstra algorithm to calculate the best path to a target network. The best path can be determined by one or more metrics such as hops, delay, or bandwidth. Once this path has been determined, the router will inform other routers as to its findings. This is how reliable routing tables are developed and routing tables reach convergence. Link state routing is considered more robust than distance vector routing protocols. One reason is because link state protocols have the ability to perform faster routing table updates.

NOTE

Convergence is the point at which routing tables have become synchronized. Each time a network is added or dropped, the routing tables must again resynchronize. Routing algorithms differ in the speed at which they can reach convergence.

Hacking Routers

Full control of a router can often lead to full control of the network. This is why many attackers will target routers and launch attacks against them. These attacks may focus on configuration errors, known vulnerabilities, or even weak passwords.

Router Attacks

Routers can be attacked by either gaining access to the router and changing the configuration file, launching DoS attacks, flooding the bandwidth, or routing table poisoning. These attacks can be either hit-and-run or persistent. Denial of Service attacks are targeted at routers. If an attacker can force a router to stop forwarding packets, then all hosts behind the router are effectively disabled.

Router Attack Topology

The router attack topology is the same as all attack topologies. The steps include:

1. Reconnaissance
2. Scanning and enumeration
3. Gaining access
4. Escalation of privilege
5. Maintaining access
6. Covering tracks and placing backdoors

Tools & Traps…

Hardening Routers

The Router Audit Tool can be used to harden routers. Once downloaded, RAT checks them against the settings defined in the benchmark. Each configuration is examined and given a rated score that provides a raw overall score, a weighted overall score (1-10), and a list of IOS commands that will correct any identified problems.

Denial-of-Service Attacks

Denial-of-service (DoS) attacks fall into three categories:

- **Destruction.** Attacks that destroy the ability of the router to function.
- **Resource consumption.** Flooding the router with many open connections simultaneously.
- **Bandwidth consumption.** Attacks that attempt to consume the bandwidth capacity of the router's network.

DoS attacks may target a user or an entire organization and can affect the availability of target systems or the entire network. The impact of DoS is the disruption of normal operations and the disruption of normal communications. It's much easier for an attacker to accomplish this than it is to gain access to the network in most instances. Smurf is an example of a common DoS attack. Smurf exploits the Internet Control Message Protocol (ICMP) protocol by sending a spoofed ping packet addressed to the broadcast address and has the source address listed as the victim. On a multiaccess network, many systems may possibly reply. The attack results in the victim being flooded in ping responses. Another example of a DoS attack is a SYN flood. A SYN flood disrupts Transmission Control Protocol (TCP) by sending a large number of fake packets with the SYN flag set. This large number of half-open TCP connections fills the buffer on victim's system and prevents it from accepting legitimate connections. Systems connected to the Internet that provide services such as HTTP or SMTP are particular vulnerable.

DDoS attacks are the second type of DoS attack and are considered multiprotocol attacks. DDoS attacks use ICMP, UDP, and TCP packets. One of the distinct differences between DoS and DDoS is that a DDoS attack consists of two distinct phases. First, during the preattack, the hacker must compromise computers scattered across the Internet and load software on these clients to aid in the attack. Targets for such an attack include broadband users, home users, poorly configured networks, colleges and universities. Script kiddies from around the world can spend countless hours scanning for the poorly protected systems. Once this step is completed the second step can commence. The second step is the actual attack. At this point the attacker instructs the masters to communicate to the zombies to launch the attack. ICMP and UDP packets can easily be blocked at the router, but TCP packets are difficult to mitigate. TCP-based DoS attacks comes in two forms:

- **Connection-oriented.** These attacks complete the 3-way handshake to establish a connection. Source IP address can be determined here.
- **Connectionless.** These packets SYN are difficult t trace because source

An example of a DDOS tool is Tribal Flood Network (TFN). TFN was the first publicly available UNIX-based DDoS tool. TFN can launch ICMP, Smurf, UDP, and SYN flood

attacks. The master uses UDP port 31335 and TCP port 27665. TFN was followed by more advanced DDoS attacks such as Trinoo. Closely related to TFN, this DDoS allows a user to launch a coordinated UDP flood to the victim's computer, which gets overloaded with traffic. A typical Trinoo attack team includes just a few servers and a large number of client computers on which the Trinoo daemon is running. Trinoo is easy for an attacker to use and is very powerful in that one computer is instructing many Trinoo servers to launch a DoS attack against a particular computer.

Routing Table Poisoning

Routers running RIPv1 are particularly vulnerable to routing table poisoning attacks. This type of attack sends fake routing updates or modifies genuine route update packets to other nodes with which the attacker attempts to cause a denial of service. Routing table poisoning may cause a complete denial of service or result in suboptimal routing, or congestion in portions of the network.

Hit-and-Run Attacks and Persistent Attacks

Attackers can launch one of two types of attacks, either-hit and-run or persistent. A hit-and-run attack is hard to detect and isolate as the attacker injects only one or a few malformed packets. With this approach, the attacker must craft the attacks so that the results have some lasting damaging effect. A persistent attack increases the possibility for identification of the attacker as there is an ongoing stream of packets to analyze. However this attack lowers the level of complexity needed by the attacker as they can use much less sophisticated attacks. Link state routing protocols such as OSPF are more resilient to routing attacks than RIP.

Damage & Defense…

Forensic Analysis of Routing Attacks

During a forensic investigation the analyst should examine log files for evidence such as IP address and the protocol. It is a good idea to redirect logs to the syslog server. This can be accomplished as follows:

```
#config terminal
Logging 192.168.1.1
```

Investigating Routers

When investigating routers there are a series of built-in commands that can be used for analysis. It is unadvisable to reset the router as this may destroy evidence that was created by the attacker. The following show commands can be used to gather basic information and record hacker activity:

- Show access list
- Show clock
- Show ip route
- Show startup configuration
- Show users
- Show version

Chain of Custody

The chain of custody is used to prove the integrity of evidence. The chain of custody should be able to answer the following questions:

- Who collected the evidence?
- How and where is the evidence stored?
- Who took possession of the evidence?
- How was the evidence stored and how was it protected during storage?
- Who took the evidence out of storage and why?

There is no such thing as too much documentation. One good approach is to have two people work on a case. While one person performs the computer analysis, the other documents these actions. At the beginning of an investigation, a forensic analyst should prepare a log to document the systematic process of the investigation. This is required to establish the chain of custody. This chain of custody will document how the evidence is handled, how it is protected, what process is used to verify it remains unchanged, and how it is duplicated. Next, the log must address how the media is examined, what actions are taken, and what tools are used. Automated tools such as EnCase and The Forensic Toolkit compile much of this information for the investigator.

Volatility of Evidence

When responding to a network attack, obtaining volatile data should be collected as soon as possible. Although all routers are different, you will most likely be working with Cisco products as Cisco has the majority of the market share. Cisco routers store the current

configuration in nonvolatile ram (NVRAM). The current configuration is considered volatile data and the data is kept in Random Access Memory (RAM). If the configuration is erased or the router powered down all information is lost. Routers typically are used as a beachhead for an attack. This means the router may play an active part in the intrusion. The attacker uses the router as a jumping off point to other network equipment.

When starting an investigation you should always move from most volatile to least volatile. The first step is to retrieve RAM and NVRAM. To accomplish this you may use a direct connection to the console port using RJ-45-RJ-45 rolled cable and an RJ-45-to-DB-9 female DTE adapter. In instances when a direct connection is not available a remoter session is the next preferred method. Insecure protocols such as FTP should not be used; an encrypted protocol Secure Shell (SSH) is preferred. You should make sure to capture both volatile and nonvolatile configuration for comparison changes and documentation purposes. Cisco routers have multiple modes, so to gain privilege mode the password must be known by the analyst.

Case Reports

Case reporting is one of the most important aspects of computer forensics. Just as with traditional forensics everything should be documented. Reporting should begin the minute you are assigned to a case. Although it may sometimes seem easier to blindly push forward, the failure to document can result in poorly written reports that will not withstand legal scrutiny.

Let's face it, not all aspects of computer forensics are exciting and fun. Most of us view paperwork as drudgery. It is a somewhat tedious process that requires an eye for detail. Don't allow yourself this fallacy. In the end, the documentation you keep and the process you follow will either validate or negate the evidence. The report is key in bringing together the three primary pieces of forensics: acquisition, authentication, and analysis.

The case report will be the key to determining one of the following actions:

- Employee remediation
- Employee termination
- Civil proceedings
- Criminal prosecution

When the investigation is complete a final written report is prepared. Some of the items found in this report will include:

- Case Summary
- Case Audit Files
- Bookmarks

- Selected Graphics
- File Location Path
- File Location Properties

Although this is not an all-inclusive list it should give you some indication of what should be included. Depending on the agency or corporation, the contents of the report will vary. What is consistent is that anyone should be able to use the logs and the report to recreate the steps performed throughout the investigation. This process of duplication should lead to identical results.

Incident Response

Incident response is the effort of an organization to define and document the nature and scope of a computer security incident. Incident response can be broken into three broad categories that include:

- **Triage.** Notification and identification
- **Action/Reaction.** Containment, analysis, tracking
- **Follow up.** Repair and recovery, prevention

Compromises

Before a compromise can be determined, investigators must be alerted that something has happened. It is best if the alert function is automated as much as possible. Otherwise, the sheer volume of log information would be overwhelming for an employee. Even with a high level of automation someone must still make a judgment regarding the validity of the alert. Once an attack has been validated it is important to reduce the damage of the attack as quickly as possible and work to restore normal business functions.

Summary

In this chapter, we reviewed how routers can play an important part in forensics. Readers were introduced to routed protocols such as IP and we discussed how routed protocols work. In many ways, IP acts as a "postman" since its job is to make the best effort at delivery. In a small network or those that seldom change, the route that the IP datagrams take through the network may remain static or unchanged. Larger networks use dynamic routing. Administrators use routing protocols such as RIP for dynamic routing. We also looked at how attackers attack routers and how incident response relates to routers and router compromises.

Frequently Asked Questions

Q: Why is network forensics important?

A: Network forensics are important because so many common attacks entail some type of misuse of network resources.

Q: What are the different ways in which the network can be attacked?

A: Attacks typically target availability, confidentiality, and integrity. Loss of any one of these items constitutes a security breach.

Q: Where is the best place to search for information?

A: Information can be found by either doing a live analysis of the network, analyzing IDS information, or examining logs that can be found in routers and servers.

Q: How does a forensic analyst know how deeply to look for information?

A: Some amount of information can be derived from looking at the skill level of the attacker. Attackers with little skill are much less likely to use advanced hiding techniques.

Q: Where do routers reside in relationship to the OSI model?

A: Routers are a layer 3 device.

Q: Do routers pass physical addresses?

A: No, not by default since routers are layer 3 devices and physical addresses are found at layer 2.

Q: What do routers do with broadcast traffic?

A: Routers block physical broadcast traffic.

Q: Why target routers?

A: Routers can sometimes be overlooked by security professionals since so much time is placed on securing workstations and servers.

Q: What is the first thing an attacker does when targeting a router?

A: An attacker must first identify the device and be able to verify it is a router. With this done the attacker must next determine the version and model of the router.

Q: What is the most important preplanning aspect of router forensics?

A: You must make sure good policies and procedures are in place that specify adequate logging is taking place.

Q: What type of skills are required for incident response?

A: Incident response requires technical skills, investigative skills, and leadership skills.

Q: How would you best define the incident response process?

A: Incident response is the process of detecting a problem, determining its cause, minimizing the damage.

Chapter 16

Investigating Wireless Attacks

Solutions in this chapter:

- **Basics of Wireless**
- **Wireless Penetration Testing**

- ☑ **Summary**
- ☑ **Frequently Asked Questions**

Introduction

In many corporate networks across the country, wireless devices are extending the network perimeter beyond the confines of the office walls and into neighboring buildings and public streets. No longer does an attacker need to gain network access by breaking into an office or attempting to bypass strict firewall policies. Now they can take advantage of an unsuspecting company who doesn't realize their wireless infrastructure security is so loose it can be hacked in less than 15 minutes.

You often read stories about hackers sitting in vehicles with a laptop and a high-powered wireless antenna, scanning for insecure wireless networks to compromise. These hackers may be looking for internal corporate intellectual property to sell to a competitor or data that may be used for blackmail or extortion. The more difficult you make it for the attacker, the better chance you have that a hacker will pass on your network and go after a less secure one.

Wireless attacks are occurring! BJ Wholesale Club, Lowe's Companies Inc., DSW, Wake Forest University School of Medicine, and TJX have all been victims of compromised wireless networks. The cost of these data breaches could be well into the millions of dollars, but their true cost remains difficult to calculate since we must take into consideration the money spent by the company to investigate the problem, provide damages to the victims, fix the vulnerabilities, and factor in loss of revenue due to low consumer confidence.

Regardless of the complexity of your design, an attacker may still attempt to compromise your wireless network. It is no longer a matter of "if" your wireless network will be attacked, but "when." External attempts to connect to your wireless network can range from simple automatic device associations to specific attacks to gain access. It can be difficult to determine whether someone is attempting to attack your network or looking for free Internet access. Regardless of their intent, every effort should be made to make your wireless network as secure as possible. Ensuring you configure strong encryption and adhere to secure wireless best practices will help minimize your risk.

As the cost to implement wireless networks continues to decrease, more and more companies are considering deploying wireless technology. Since 2006, London has seen 160 percent more wireless access points, with New York increasing a substantial 49 percent. However, you don't need to look at statistics to realize that wireless networks are becoming more popular. Open up your wireless configuration user interface and you will usually see three or four networks in the surrounding area. If you live in the city, that number can be as high as 15 to 20.

Deploying a wireless network has numerous advantages. Now, workers can access their corporate network anywhere, from the cafeteria to the conference room, without having to reconfigure their computers or plug into a wall jack. This flexibility is not only beneficial for the employee, but the employer as well. It is often much cheaper to implement a wireless solution than installing network cabling throughout the building.

This chapter discusses the basics of wireless networks and the methods hackers use to attack them. It is imperative you understand the techniques used by hackers so you can test your wireless networks for vulnerabilities. Hackers have an advantage when it comes to wireless attacks due to anonymity and the difficulty of tracking down attacks. The attack surface is huge and the results are appealing because once a hacker gets in, they are on the internal network, bypassing many of the traditional network security barriers.

Basics of Wireless

A wireless access point connects the traditional wired network to wireless clients by transmitting network data through the air. The wireless access point acts as a relay between the wired network and wireless clients. Many companies implement many wireless access points that can hand off the wireless signal to ensure the client does not get disconnected. These access points range from home or home office devices that cost less than a hundred dollars to enterprise access points that implement complex security features and allow for specialized configuration and management.

The wireless access points themselves simply provide a means for the data to travel from the wired network to the client. The network infrastructure services must still be provided to the client. While most access points today have these features built into the device, you can also use your traditional servers to provide these services such as DHCP and routing.

Different types of wireless networks are available, the most common being 802.11 b/g. 802.11 b was the first wireless standard developed for commercial use. It operates at 2.4 GHz and transfers data at a rate of 11 megabits per second (Mbps). Improving data transfer to 54 Mbps, 802.11 g was a welcome upgrade for consumers wanting faster throughput from their wireless infrastructure. Although the 2.4 GHz band would often lead to interference problems, backwards compatibility and early adoption helped this combination become the standard for wireless devices.

For those companies that required less interference or that could not adopt the 2.4 GHz band as an acceptable solution, 802.11a was released that worked on the 5 GHz band. 802.11a improved on wireless stability, reducing dropped connections, while staying at the 54-Mbps transfer rate. The main downside is the expense of upgrading both the wireless access points and the client's wireless cards to devices that support this standard.

Wireless devices can also connect two remote locations by using line of sight antennas to send data from one location to another where the installation of cabling would not be feasible. A company who has a second building located less than a mile from the main building may choose to implement a line of sight wireless solution for Internet connectivity instead of spending the reoccurring cost of a traditional T1 line.

For a client to connect wirelessly to a LAN, they must have a wireless network card compatible with the wireless standard used by the target network. With the majority of laptops today including built-in network devices, wireless networks are becoming a standard

with most home and corporate networks. While wireless LANs have traditionally been used for mobile devices, the ease and cost effectiveness of extending the network has some companies implementing wireless cards on their desktop machines.

Even with all the advancements in recent years, installing a wireless network on your LAN is not something you should do without careful planning. Before adding wireless connectivity to your network, you should understand the major advantages and disadvantages of introducing wireless connectivity.

Advantages of a Wireless Network

Wireless networks offer users the following advantages:

- **Ease of accessing the network** No longer are employees bound to areas where they can plug into a wall jack to connect to the network. Being able to connect your laptop to the network from anywhere in the building allows for more freedom.
- **Reduce cost of running cable** The labor cost alone of wiring an office or office building can be quite high compared to the cost of purchasing and installing a wireless device and cards. With wireless cards available for desktop computers, an office can be network available in a manner of minutes.
- **Productivity** Employees can now take their laptops with them to meetings or any other place required, and yet continue working.

Disadvantages of a Wireless Network

Wireless networks also have some disadvantages as well:

- **Security** The main disadvantage of wireless is its lack of security. If proper planning and configuration are not taken into consideration, any wireless network can be easily compromised.
- **Complexity and reliability** While adding wireless access points increases accessibility to the wired network, it can also increase the complexity of the network design. System administrators must be aware of the impact that the wireless devices have on the network and troubleshoot problems such as weak signals and dropped wireless connections.
- **Network performance** While the advertised speed of a wireless connection is 54 Mbps, the actual data throughput is often much less, especially if many computers are using the same wireless connection. A wired network drop will almost always outperform a wireless connection.

Association of Wireless AP and a Device

To transmit and receive data through a wireless connection, the client must be associated with the access point. This association connects the two devices and allows the client to obtain an IP address and communication on the network. Signal power and security settings can interfere with the association process.

On Windows XP and Windows Server operating systems, Microsoft provides the Windows Wireless Zero Configuration (WZC) utility to assist the user with connecting to wireless networks. While it does make the association process easier, it also introduces some security concerns you should be aware of.

If a preferred network is not available, WZC will probe for networks that are already connected. This information can be viewed by anyone using a wireless analyzer and can be used to set up fake access points to lure clients to connect. WZC will also attempt to connect to the wireless network with the strongest signal. Knowing this, an attacker can create fake wireless networks with high-power antennas and cause computers to associate with their access point rather than the legitimate one. When possible, it is recommended you use the vendor-provided wireless management tool.

If you are using a wireless card from Linksys, Dell, or Netgear, you will likely have installed management software from the vendor that can control the association of the wireless device. This software has been designed by the manufacturer to work with the corresponding hardware and is typically more secure.

All wireless management software will require you to select the wireless network to which you would like to connect. You will select the Service Set Identifier (SSID). The SSID is the public name of the wireless network. If the network has security enabled, you will be prompted to enter the passphrase or encryption key. If you do not know the key, you will not be allowed access to the network. If the network is open or you enter in the correct security key, your computer will attempt to establish the connection.

Access Control

To secure the wireless network as much as possible, security controls can be implemented to minimize the ability of an attacker to gain access. These access controls can be used independently or combined to provide additional protection. The access controls available to you will depend on the type of wireless access point you are configuring. Most access points will allow the configuration of encryption and Media Access Control (MAC) filtering.

Encryption

Wireless data is being transmitted in the air that is susceptible to nonauthorized users capturing this data and reading the contents. To eliminate data being sent over the air in cleartext, wireless networks adopted encryption as part of the network standard. Unfortunately, early

implementations of encryption for wireless soon became trivial to crack. As wireless networking matures, the encryption is becoming more complex and difficult to crack. Different technology is also being implemented, such as certificate-based encryption, as an alternative to preshared keys.

WEP

Wired Equivalent Privacy (WEP) is an optional security feature that was specified by the 802.11 protocol to provide authentication and confidentiality in a wireless access point. It was one of the original methods of sending secure information using a wireless network. When the IEEE committee recommended WEP be used as a mechanism, it also said WEP should not be considered an adequate security and strongly recommended it not be used without an authentication process for key management. WEP employs a symmetric key to authenticate wireless devices and guarantee the integrity of the data by encrypting the transmissions. Each wireless access point and client must share the same key in order for authentication to take place. Once WEP has been enabled, it then begins a challenge and response authentication process. WEP encrypts data before it is sent from the system, and then decrypts it at the access point. It has been shown to be severely inadequate in its security methods, and as such was replaced as the wireless encryption method of choice in 2003. It uses the stream cipher RC4 (Rivest Cipher). Its method of authentication is essentially that of Shared Key Authentication since both the access point and the wireless device possess the same key. WEP keys are always either 40 bit or 104 bit. The advertised 64-bit and 128-bit encryption results from the fact that the Initialization Vector (IV) is 24 bits. With the encryption method used by WEP, an attacker with enough IVs can crack the key used and gain full access to the network. In studies, WEP networks have been broken into extremely quickly. Recently, at an information security conference, the Federal Bureau of Investigation demonstrated how to crack WEP—and they did it in three minutes.

Overall, the use of WEP can create in the casual user a false sense of security. In the worst-case scenario, an attacker can gain full access to the network extremely quickly—as such, this should not be used to secure a sensitive network. In today's access points, WEP is not considered a secured method of securing a wireless access point.

WPA

WPA resolves the issue of weak WEP headers as explained earlier, or the IVs, and provides a way of insuring the integrity of the messages that passed the integrity check using TKIP (Temporal Key Integrity Protocol) to enhance data encryption. WPA-PSK is a special mode of WPA for home users without an enterprise authentication server and provides the same strong encryption protection. In simple terms, WPA-PSK is extra-strong encryption where encryption keys are automatically changed (called rekeying) and authenticated between devices after a specified period of time, or after a specified number of packets has

been transmitted. This is called the rekey interval. WPA-PSK is far superior to WEP and provides stronger protection for the home/SOHO user for two reasons. The process used to generate the encryption key is very rigorous, and the rekeying (or key changing) is done very quickly.

WPA improves over WEP because it uses a per-session encryption key. Every time a station associates, a new encryption key is generated based on randomization and the MAC addresses of the wireless access point.

Unfortunately, though, the easiest way to use WPA actually makes it easier to crack than WEP. When 802.1X authentication is not used in WPA, a simpler system called Pre-Shared Key (PSK) is utilized instead. A pre-shared key is a password that all clients need to be configured with in order to access the access point. Most consumer routers have the capability of using the WPA with the PSK pre-shared key.

With WPA-PSK, and almost all passwords, if you make a short character password, you are susceptible to an offline dictionary attack where an attacker grabs a few packets at the time a legitimate station joins the wireless network, and then can take those packets and recover the PSK used. An attacker can get what he needs in order to guess the PSK and get out without anyone noticing. This can occur because the attacker doesn't have to be near the WLAN for more than a few seconds, and the LAN doesn't have to be very busy. This attack depends on the choice of a password, although cracking techniques get better and better. Thus, WPA has been defeated. Most wireless access points have a mechanism built into WPA, which converts an 8- to 63-character string you type in to the 64-digit or 128-digit key (as used with WEP). However, most wireless access points won't be able to use the whole 64-bit key with passphrase mode.

The innate problem is that a passphrase is easy to guess. The IEEE committee that wrote 802.11i pointed out that an eight- to ten-character passphrase actually has less than the 40 bits of security that the most basic version of WEP offers, and said that a passphrase of less than 20 characters is unlikely to deter attacks. As with WEP, wireless cracking tools exist that are specifically designed to recover the PSK from a WPA-protected network (such as Kismet) which are easily available to download.

WPA with 802.1X authentication (sometimes called WPA-Enterprise) makes for a much more secure network. 802.1X offers strong positive authentication for both the station and the WLAN infrastructure, while deriving a secure, per-session encryption key that is not vulnerable to any casual attack. This is typically used with a RADIUS server as explained earlier regarding authentication. The best wireless security mechanism out there with the most access points is 802.1X authentication, combined with WPA's improved encryption.

But instead of pointing out all of the inherent flaws in passing data over unlicensed radio frequencies, WPA, which comes with most new consumer routers, is an excellent way to keep your Internet surfing and home network as safe as it can be. Add VPN connections and MAC filtering and you basically have the same security as you do with an alarm system on your home. It deters people from getting access without considerable resolve.

MAC Filtering

With few exceptions, network devices contain a physically unique burned-in MAC address. The purpose of this is to give a unique identifier to that piece of equipment. It is pre-assigned to the devices by its manufacturer, and in theory is completely unique. In most cases, the MAC address is 48 bits long. A standard format exists for writing MAC addresses, which consist of three groups of four hexadecimal digits separated by dots. The more common method of writing is to use six groups of two hexadecimal digits, which are in turn separated by colons or hyphens—for example, 00-07-E9-E3-84-F9.

Since each MAC address is unique, it can be used to limit network access. The steps to do so will vary by access point, but will always involve the following:

- Finding the MAC address for the devices that will be allowed to access the network (this can be found either by looking on the actual device itself, or by using the *ipconfig /all* command in the Windows command console, and *ifconfig – a* for Linux and OSX consoles).
- Enter the MAC address into the configuration for the access point (this will vary by device).

In theory, once the MAC addresses are entered into the access point, they will be the only devices allowed to access the network. In practice, a number of issues can come up, including, but not limited to, the fact that any time a new system is used in conjunction with the access point (such as a visiting client who needs to connect to the network), its MAC address must be entered individually into the access point. This can pull the network administrator away from other tasks of monitoring the network. The administrator must also periodically remove the information for devices no longer being used.

Another problem that exists is that of MAC address spoofing. While the address is typically encoded to the physical medium of the network device, software exists that can make a device access point have a different MAC address than it actually does. While this does have legitimate and useful purposes, including privacy and interoperability, it can also be changed to access a system illegally. Because of this large security hole, MAC address filtering should not be used by itself, and should instead be included in a wider security policy involving encryption and other authentication methods.

Cloaking the SSID

The SSID is essentially the name of the access point or wireless network seen by the user. A client can find out their SSID in only two ways: it can be told to them by the access point actively, or you can passively put it in the client's configuration. When it is given out automatically, it is called Open Network mode, but is known as Closed Network mode when "cloaking," meaning the SSID is not broadcast over the radio frequency. The automatic

giving out of the SSID occurs typically every 100 milliseconds and is called a beacon, within which is contained synchronization information such as channel, speeds, timestamps, encryption status, and other information.

In Closed Network mode, the SSID is not broadcast out to the user or administrator programs. Thus, the client must probe the access point, and if the SSID matches, it will synchronize and go through the authentication process. Authentication can occur using an open system or shared key authentication. An open system does not need any credentials supplied. The SSID is used in wireless networks to identify the wireless access point and its associated network. It can be up to 32 alphanumeric characters long and is attached to all packets sent over the wireless connection. Additional access points in the area can broadcast the same SSID, so its use as a security or authentication method is negligible. However, the SSID may be changed and cloaked (meaning it is not set to be broadcast by the AP) in an attempt to minimize the visibility of the network. By default, most APs broadcast their SSID to the surrounding area. This beacon method occurs roughly every 0.1 seconds. Most APs ship with a default SSID which are widely known and can be found on a number of Internet sites. Using a default SSID may draw malicious users who suspect that additional settings on the AP are set to default settings as well (such as the administrative password). From a security standpoint, changing the SSID may cause some possible attackers to choose a wireless network that uses default settings. Changing the SSID is reasonably simple for a nontechnical user, and though minor, is a step toward securing the AP. It is important to note that disabling the broadcast of the SSID does not completely cloak the AP. It simply makes it less visible.

Notes from the Underground…

Wigle (Wireless Geographic Logging Engine)

Wigle (www.wigle.net) is the consolidation of thousands of wardriving efforts to create a database of wireless networks around the world. Wigle accepts WAP information by NetStumbler, DStumbler, Kismet, and Pocket Warrior exports. You can also use the Web site's online submission form to enter wireless networks.

In just over five years, Wigle has collected information on 10,953,561 unique wireless networks with locations. Of these networks, 44.4 percent are encrypted with WEP (which we know can be cracked in minutes) and 18.5 percent still maintain the default SSID or network name, which leads us to believe these networks may still maintain the vendor's default settings.

The danger of so much wireless access point data being available to hackers is that by searching for APs by SSID or network name, BSSID or MAC address, longitude and latitude, or street address, anyone can find an insecure wireless network.

Wireless Penetration Testing

Many managers and system administrators never bother to learn the techniques hackers use to attack wireless networks. A negative stigma is often associated with the use of hacker tools and methodologies. They frequently consider the use of these tools as the legitimization of hacker-based tactics and strategies. This attitude can lead to an insecure wireless network that has not been tested to determine the ability to prevent a true attack from succeeding.

Conducting penetration tests to determine the security of your wireless infrastructure is essential. Through the practice of learning, understanding, and implementing the same methods of attack the intruder will use, one can better assess vulnerabilities, overcome weaknesses, and fortify defenses. During a wireless penetration test, it is important to gather as much information about the network as possible. Most networks are discovered during wardriving (the process of scanning for insecure networks), so the attack may be targeted more towards a wireless vulnerability than a specific corporate network.

The two approaches you want to take when doing a penetration test are an internal and external attacker. The internal approach allows you to use information you already know about the network, like encryption keys, network design, and signal ranges. This type of test validates the security of the network from an internal employees point of view. External testing is done without using any knowledge of the network infrastructure. The tester uses tools an intruder would use and simulates an actual attack. It is important that a skilled security professional perform this test to ensure these tools do not negatively impact network and server systems.

The key thing to remember during a wireless penetration test is that your goal is to assess the security of the wireless network. Management approval in writing of what tests are expected and the potential impact to the network is a good way to ensure no surprises for either side.

- You should never do anything that will compromise or affect a neighboring wireless network.
- You should not try any type of Denial-of-Service attack on the network without approval.
- You will be attempting to crack the various portions of the network. Make sure to keep the result of the scan and penetration test confidential, and ensure they are stored safely.
- Report all findings to management with detailed explanations and security recommendations.

Search Warrants

Before you can investigate on a device owned by someone else, you may be required to obtain a search warrant. You must verify that the search warrant includes the proper authorization to

perform onsite examinations of computer equipment such as wireless access devices. Do not perform any forensic analysis on equipment you are not authorized to analyze.

Direct Connections to Wireless Access Point

Users that are persistent in their desire to connect to an unauthorized access point will keep an eye out for security professionals completing wireless audits and simply unplug the access point until the scan is complete. The amount of effort required for physical wireless scans can often result in less frequent scans. Detecting wireless access points from your wired network has many advantages. You can set automated scripts that search your network on a continuous basis, saving you time and money. By using tools like Network Mapper (Nmap), you can scan sections of your network not easily accessible to wireless scanning.

To perform wired network scans for access points, you must be connected to the internal network and have the ability to connect to all the subnets you want to scan.

Scanning for Wireless Access Points with Nmap

One of the more popular tools for completing network scans is the Network Mapper (Nmap), a Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) Internet Protocol (IP) scanner. Nmap supports several different scanning styles, the most important being "stealth" scanning. "Flying under the radar" of the target system's administrator is crucial to the intruder's successful attack, and stealth scanning has the advantage of being able to pass through most firewall and network monitoring systems unmolested and largely unnoticed.

Using these scans, we can determine what ports are open on our network systems and identify unauthorized wireless access points. Nmap is used to identify the OS of the target system, but beware, OS fingerprinting scans are easily detectable and will be detected right away by Intrusion Detection Systems (IDSs).

Nmap is a free open-source utility for network exploration and security auditing. It was designed to rapidly scan large networks, although it works fine against single hosts. Nmap uses raw IP packets in novel ways to determine what hosts are available on the network, what services (application name and version) those hosts are offering, what operating systems (and OS versions) they are running, what type of packet filters/firewalls are in use, and dozens of other characteristics. Nmap runs on most types of computers and both console and graphical versions are available. Nmap is free and open source.

The text description, vendor name, operating system, and device type provided by Nmap during OS fingerprinting is what we will use to find wireless access points (WAPs) by scanning from the wired side of the network. Nmap fingerprinting is very effective in determining WAP devices. To determine if an access point was connected to the network, we can run a scan of a local class c subnet using the following command:

```
nmap –A –T4 192.168.0.1/24
```

This command does a few things. The –A tells Nmap to enable operating system and version detection. The –T4 sets the timing template to 4, which speeds up the scan by setting the timeout value to 500 milliseconds. This can be extremely useful when scanning large subnets to ensure speedy results. What we are looking for in the Nmap output is Device type: WAP. This is our verification that we have scanned a wireless access point on the network.

Scanning for Wireless Access Points with Nessus

Nessus is a powerful and up-to-date open source scanner that scans the network for systems with security vulnerabilities. Nessus is very fast, reliable, and has a modular architecture that allows you to fit it to your needs. Scans can be tailored to seek out only those vulnerabilities you deem important. Each security test is written as an external plug-in. This way, you can easily add your own test without having to read the code of the Nessus engine. The Nessus scanner is made up of two parts: a server, which performs the security tests, and a client that serves as the front end. You can run the server and the client on different systems. Additionally, several clients are available: one for X11, one for Win32, and one written in Java.

Nessus plugin #11026 was written to detect the presence of wireless access points on the network using a variety of techniques. It scans the device via the TCP/IP Nmap fingerprint, analyzes the HTTP management interface, and verifies the FTP banner and SNMP information present. If any of these techniques determines that the device is a wireless access point, the scan moves on to the next device, marking this device as a WAP.

To scan using the #11026 plug-in for Nessus, complete the following steps:

- Run nessus-update-plugins to ensure the plug-ins are up-to-date.
- Start a new scan and ensure that the Access Point Detection plug-in is selected from the General plug-in section in the Plugins tab of the client.
- Make sure that the Enable Dependencies At Runtime option is checked since the plug-in requires the system info type via SNMP and HTTP Server–type dependencies to be able to scan.

Rogue Access Points

For enterprise-class access point detection, you may want to consider investing in a wireless IDS/IPS system. A wireless IDS/IPS is an Intrusion Detection/Intrusion Prevention System that provides 24/7 monitoring on the network and can dynamically respond to wireless threats.

In addition to detecting rogue devices, a wireless IDS/IPS can terminate them using air or port suppression on the switch, provide forensic analysis of packets sent and received, and provide location tracking of the device by triangulating the signal between multiple sensors.

Rogue access points represent a growing threat to network administrators because users can plug in a wireless access point with a minimal configuration and extend the network wirelessly. This creates a lack of control on where data is being sent, and worse, who is listening.

Wireless Connect to a Wireless Access Point

In this section, we are going to discuss the techniques hackers use to compromise wireless access points. Before initiating an attack, a hacker will want to know as much about the target network as possible, using various wireless tools to perform information gathering. Once the hacker feels comfortable with the information gathered, they may use injection techniques to force information across the wireless network that can be used to crack encryption schemes. Once these data have been captured, it is matter of time before the encryption key is discovered and the hacker is connected to the internal network via the wireless access point.

The tools used to perform these attacks are freely available on the Internet and included in live security distributions. We will be using many of the utilities from the Aircrack-ng suite (www.aircrack-ng.org) along with Kismet (www.kismetwireless.net) a popular wireless analyzer. While you may be able to find tools for other Operating Systems, most are written for Linux, so some knowledge of Unix commands will be helpful.

Notes from the Underground…

BackTrack

Getting all the tools necessary to perform penetration-testing tasks can be quite daunting. Compatibility issues with kernels and dependent drivers can have a security analyst spending most of the time getting the tools to work instead of testing the security of their applications or devices. Live CDs are bootable CDs that contain preconfigured operating systems and tools that are distributed by groups over the Internet. One of the premiere Security distributions is BackTrack (www.remote-exploit.org/backtrack.html).

The BackTrack live CD distribution comes with many security tools already installed and ready to use so that the tester can spend more time performing the penetration testing and finding security vulnerabilities in the systems. BackTrack not only contains the majority of the tools that will be discussed in this chapter, but also contains tools for Information Gathering, Network Mapping, Vulnerability Identification, Penetration Testing, Privilege Escalations, VOIP and Wireless analysis, and Digital Forensics. All you need to get started is to download the ISO image and burn it to a CD. See Chapter 11 for a detailed discussion of BackTrack.

Information Gathering

The first thing a hacker will do when trying to gain unauthorized access to a wireless network is to do some reconnaissance and get more information on what tools will be needed next. The goal of this step is to detect rogue wireless access points, ad hoc wireless networks, and open or improperly configured access points that can be used to gain access. Hackers will want to get as much information about the wireless access points as they do the associated clients because all this information will be used to gather important connection data.

Kismet

Time to see if our network is available. To do this, we will start by using Kismet. Kismet is a layer 2 wireless packet analyzer that works with wireless cards that support raw monitoring mode. Kismet can quickly identify wireless networks by capturing wireless packets and retrieving information such as whether it has security-enabled SSID cloaking, network channels, and an SSID of the network.

This step will simulate what an attacker will look for when they are scanning for vulnerable networks. After starting Kismet, you will start to see wireless networks populate the screen. By default, Kismet sorts the network lists by Auto-fit, which can be difficult to read if there are a lot of networks available. Our first step is to sort the access points by selecting "s" to get to the sort menu, and selecting "s" again to sort by SSID. This view not only sorts the network list by SSID, but groups them into categories such as Probe Networks and Ad-hoc Networks. This will be important to know as we continue to gather information about our target wireless network. To expand/contract a grouping, highlight the group and press the **Spacebar**.

Now that we have our Network List screen sorted by SSID, let's go through the different columns.

- Name: SSID of network
- T: Type of network (A =Access Point, H = Ad-Hoc, P = Probe request, A = wireless client searching for a network, D = Data network, T = Turbocell network, and G = Group of wireless networks)
- W: Identifies if network is secured (Y = Yes, N = No)
- Ch: Channel number of network
- Packts: Number of packets captured
- Flags: Method in which IP was gathered
- IP Range: IP of the network

The Kismet colors make identifying networks easier. The following are the possible color combinations:

- Yellow: Unencrypted network
- Red: Network is using factory defaults
- Green: Secured Network
- Blue: Hidden networks that are cloaking the SSID

The information we will use in the next step from Kismet is the network channel and whether or not the network is encrypted. It will also be helpful to document if Kismet was able to determine an IP range or if the network was Red or Yellow. Both indicate that security is not configured adequately and that the network can be compromised.

Aircrack-ng

Aircrack-ng (www.aircrack-ng.org) is a suite of tools for auditing wireless networks. We will be using the *airodump-ng*, *aireplay-ng*, *aircrack-ng*, and *airdecap-ng* tools from the Aircrack-ng suite.

- *Airodump-ng* captures raw 802.11 packets to be used with *aircrack-ng*. *Airodump-ng* is also capable of logging the coordinates of access points.
- *Aireplay-ng* is primarily used to inject frames into wireless traffic, which will later be used by *aircrack-ng* to crack WEP and WPA-PSK keys. *Aireplay-ng* supports deauthentications, fake authentications, interactive packet replay and ARP request (re)injections.
- *Aircrack-ng* can recover keys once enough data packets have been captured. Optimizations to the standard attack algorithms make wireless encryption cracking with Airocrack-ng much faster compared to other WEP cracking tools.
- *Airdecap-ng* is used to decrypt encrypted capture files. It can also be used to strip wireless headers from capture files.

Our first command will begin sniffing the wireless packets using the *Airodump-ng* utility. The packets will be captured and written to a file that will later be used to crack the encryption key. From a command prompt, run the following command:

```
airodump-ng –w output –c 6 ath1
```

This command runs the *Airodump-ng* command and sets the capture file to output using the –w switch. Since we know from our Kismet scan the wireless network of interest is on channel 6, we use the –c switch to ensure that *airodump-ng* stays on that channel and captures as much data as possible. Finally, we tell the command to use the interface ath0. The interface may vary, depending on the wireless card you are using.

You should now see the Airodump-ng screen. The screen is broken into two sections: the top section shows the wireless access points; the bottom section shows the wireless clients.

Now that we have *Airodump-ng* running, we will look for associated wireless clients. In the top section, find the BSSID (MAC address) of your target network. Now look in the bottom section to see which clients are associated. We will need to document the BSSID and Station address to perform packet injection.

Injection

Wireless Packet injection allows the hacker to modify packets in the air so the wireless devices are forced to generate traffic that will be captured and used to crack the encryption key. If there is a lot of traffic on the network, the hacker will not need to perform any injections to force traffic to be generated. However, if the hacker is impatient, a quick way to generate traffic is to deauthenticate a client that is already authenticated. This deauthentication forces the client to reauthenticate, generating the handshaking packets that the hacker will use during the cracking phase.

To perform packet injection as part of our penetration test, we are going to use the *aireplay-ng* utility. This utility can perform multiple attack types. We will be testing fake authentication, ARP packet replay, and deauthentication.

First, we will associate the attacking machine with the target network. This will be done using the fake authentication method. To perform fake authentication, we will run the following command.

```
aireplay-ng -1 0 –e TestNet –a 00:0F:B5:29:8C:32 –h 06:18:4D:95:32:61 ath1
```

This command runs the *aireplay-ng* command and sets the attacktype to -1, which is fake authentication. The 0 sets the reassociation timing to 0 seconds. To set the wireless network name, the –e switch is followed by the SSID of the wireless network. Then, the –a switch is followed by the MAC address of the access point and the –h switch is followed by the MAC address of the network card used for the injection. We complete the command by adding the wireless interface name. If the command is successful, you should receive the following response:

```
Sending Authentication Request
Authentication successful
Sending Association Request
Association successful :-)
```

Your wireless network card is now associated with the target access point. The next step is to set up ARP request packet replay. This is the most effective way to generate the traffic needed to crack the encryption key. While the attack retransmits the same ARP packet, the access point responds with new IVs. These packets are the key to determining the encryption key.

We will now set up the ARP request replay using the *aireplay-ng* command. To perform this attack, we open up a new command window and run:

```
aireplay-ng -3 –b 00:0F:B5:29:8C:32 –h 00:13:CE:86:08:A6 ath1
```

This command runs the *aireplay-ng* command using the -3 attack type, standard ARP request replay. The –b switch is followed by the MAC address of the access point, and the –h switch is followed by the MAC address of the associated client. Once we initiate the following deauthentication command, the ARP request replay will start flooding the access point with the retransmitted ARP packets and create the needed IV packets.

To generate the deauthenticate attack, we will again use *aireplay-ng*. Once this process is initiated, we should see a substantial rise in traffic on our airodump-ng screen. To run the deauthenticate command, we open up a new command window and run:

```
aireplay-ng -0 5 –a 00:0F:B5:29:8C:32 –c 00:13:CE:86:08:A6 ath1
```

This command runs the *aireplay-ng* command using the -0 attack which means deauthentication. The next number sets the number of deauthentication packets to 5, setting this number to 0 sends continuous deauthentication packets. The –a switch is followed by the MAC address of the access point and the –c switch is followed by the MAC address of the client you are trying to disassociate. The final option is the wireless interface name.

Now that all the injection commands have been executed. You should see the data column for the wireless network you are testing steadily rise. You should also see that the ARP request replay is reading and retransmitting ARP packets to the access point. 40-bit WEP can be cracked with 300,000 IVs, while 128-bit WEP may require 1.5 million IVs. Once you have enough packets captured, it is time to see if the wireless encryption key is vulnerable to being cracked.

Cracking

Now that we have a capture file with the necessary data, we can test our encryption key to see if it is vulnerable to being cracked. To crack WEP keys, *aircrack-ng* uses various statistical attacks that involve brute forcing the key. Due to the complexity of WPA/WPA2 preshared keys, only a dictionary attack is available.

In a new command window, we will use the *aircrack-ng* command to attempt to crack the wireless encryption key. For a typical WEP attempt, we would use the following command.

```
aircrack-ng -m 00:11:22:33:44:55 -n 64 output.cap
```

This command runs the *aircrack-ng* command using the –m option followed by the MAC address of the access point we are trying to crack. This option is not required but helps focus the attack on our target access point. The –n switch specifies the length of the key. If you do not know the length, try both 64 or 128. Last, we will provide the capture file where the data captured from *airodump-ng* is stored.

If the command is successful, it will display the encryption key of the target network.

Passive and Active Sniffing

If you have discovered the encryption key of the target network, you will be able to associate your computer with the target network. At this point, your network has been compromised and it is a good time to reevaluate your wireless network's security controls. It is quite possible that hackers have already done what you have just done and have been trying to obtain data from your corporate network.

The first reaction most administrators have when they realize their network may be compromised is to check the logs to see if anyone who is not authorized has connected to the network. While this is a good first step when investigating whether you are actively connected, you must remember one thing, wireless data are being transmitted in the air and now that the attacker has the encryption key, they can sniff the network traffic and decrypt it passively. This means you will never see them connect to the network! To demonstrate this, let's take a look at the capture file you saved using *airodump-ng* to use with *aircrack-ng*.

The initial purpose of this file was to be used with *aircrack-ng* to discover the encryption key. Now that we know what the key is, we can use the *airdecap-ng* utility to strip the wireless headers and decrypt any encrypted packets, leaving us with a capture file that can read data from the target network in plain text with any packet analyzer, such as Wireshark.

To run the *airedecap-ng* command, we will open up a new command window and run:

```
airdecap-ng –e TestNet –w 145AA34FA1 output.cap
```

This command runs the *airdecap-ng* command and supplies the following options. The –e switch is followed by the ESSID of the target network. The –w switch is followed by the hexadecimal WEP key. Then the capture file is provided. This will decrypt the capture file and save it to a file name output-dec.cap. If the command is successful, you should see a similar result:

```
Total number of packets read           1658828
Total number of WEP data packets       255816
Total number of WPA data packets       0
Number of plaintext data packets       30
Number of decrypted WEP packets        255816
Number of decrypted WPA packets        0
```

Now you can open up decrypted captured files and read the content of the network traffic. This is a passive method of capturing data because once the encryption key is obtained; the attacker no longer must connect to the network to obtain network data. If the system administrator did not detect the initial packet injection attack on the access point and take appropriate action, the attacker will be able to continue sniffing data undetected until the key is changed.

If an attacker is not concerned with being detected on the network, they can be more aggressive and connect to try and sniff data directly from the wireless connection. While this technique establishes a direct connection with the target network and can be logged by the

access point, unless there is a wireless IDS/IPS in place or constant scanning to determine unauthorized MAC addresses on the network, it will still be difficult to detect.

One of the many dangers associated with active scanning is the hacker's ability to use more complex tools on the internal network to gather additional information. Instead of being able to only read wirelessly transmitted data, tools like Ettercap make it possible to sniff wired network traffic from a wireless connection as long as the computers are on the same subnet.

Tools…

Ettercap

Ettercap is a multipurpose suite for man-in-the-middle attacks on LAN. It provides the ability to sniff live connections, offers content filtering on-the-fly, and many other features. It supports the active and passive examination of many protocols and includes many features for network and host analysis.

With four modes of operation (IP, MAC, ARP, and public ARP), Ettercap is an extremely flexible utility that can provide a lot of information to an attacker and is difficult to detect. An attacker that uses Ettercap can use it for password collection, OS fingerprinting, ARP poisoning, and character injection.

Logging

Most wireless access points have the ability to log traffic and connections. It is important to consider logging requirements before an incident occurs. Being able to go back to a specific time when an event is thought to have occurred gives the investigator the ability to analyze and determine whether there was indeed a wireless attack on the network.

The logging on most access points does not provide the kind of granularity required for effective logging. Wireless IDS/IPS systems can be set up to monitor and log any suspicious wireless activity using wireless sensors. These sensors detect malformed wireless packets and log the data for forensic analysis later.

Investigating wireless attacks is difficult due to the nature of the technology. This is exactly the reason attackers like to use this method to gain access to corporate networks. The best solution is to configure your wireless network securely to ensure hackers will not be able to gain access easily. If a wireless incident does occur, an investigator can use the same tools the attacker used to determine how the attack was done and decide how much information may have been exposed. Proactively reviewing event logs on servers and network devices is always a good step towards identifying attacks early and handling them appropriately.

Summary

To investigate attacks against your wireless network, it is important to understand the techniques and methods hackers will use to gain unauthorized access as well as the inherent vulnerabilities in the technology. Understanding the insecurities of WEP versus WPA-PSK, for example, will allow you to direct your investigation appropriately so you can determine the extent of an incident and how much data may have been compromised. Determining the extent of a damage an attack has on a company is difficult with factors such as resolving the vulnerability, legal responsibility, lost intellectual property, and lost revenue due to low consumer confidence all being part of the equation.

Unfortunately, once it is realized that the wireless network has been compromised, the damage may already have been done. The best way to prevent a wireless attack is to ensure the corporate wireless access point, wireless clients, and network configuration are as secure as possible. Using WPA-PSK is a good start, but adding additional security configurations such as MAC filtering, cloaking the SSID, 24/7 rogue access point monitoring, Nmap scans, wireless access point logging, and sound wireless policies are necessary to secure a network from hackers. These techniques by themselves are not going to stop a hacker from attacking your network, but combined they will make the task much more difficult.

Conducting scheduled penetration tests to verify the security of the wireless network are also recommended. Technology changes rapidly, as do the tools that hackers use to exploit it. A wireless network that was safe last year may not be as secure this year. Staying on top of the latest trends in wireless hacking will allow you to implement secure solutions quickly and effectively. For enterprise environments where wireless security is a top priority, a wireless IDS/IPS solution can provide an extra level of security and auditing.

Frequently Asked Questions

Q: Our company is very small. Why would an attacker want to gain access to our network?

A: Most hackers are going to go after easy-to-hack networks and try to gain as much information as possible. The motivation for many black hat hackers is financial gain. If your company uses the Internet to read corporate e-mail, access password-protected Web sites, or conduct B2B or B2C transactions, a hacker may attempt to disrupt the operation. Because one of the easiest ways to gain access to a corporate network is through wireless networks, it is important to make sure your wireless configurations are as secure as possible.

Q: To hack into a wireless network would take a highly experienced hacker and I don't think anyone in my neighborhood would be able to do it.

A: The Internet has made it possible for anyone that is curious to become a viable threat when it comes to hacking, no matter where they live or how much computer knowledge they have. Long gone are the days when a hacker would have to code an application from scratch to exploit vulnerabilities. Now, it is as simple as downloading the application and following instructions. BackTrack is security operating system distribution that provides many of the tools discussed in this book, which is available free of charge. These days, the tools that security professionals use to test systems for vulnerabilities are often the same ones hackers use for malicious intent.

Q: Which encryption should I use to protect my wireless network, WEP or WPA?

A: WEP networks have proven to have weak security and can be broken into extremely quickly. Recently, the Federal Bureau of Investigation demonstrated how to crack WEP—and they did it in three minutes. The use of WEP can create a false sense of security. WPA utilizing a preshared key is far superior to WEP and provides stronger protection for two reasons, the process used to generate the encryption key is very rigorous, and the key changing is done very quickly.

Q: How do I detect if a hacker has associated to my wireless access point?

A: The first step is to look at the management interface and determine if there are any machines that do not look familiar. Most management consoles will provide the client MAC address and computer name. Since this information can be spoofed, it should not be considered 100 percent reliable. Instability in the network may be another indication that an attacker is attempting to capture corporate data. The best solution is to implement an Intrusion Detection/Prevention System. This system would monitor the network and search for suspicious network traffic, providing alerts to security and system administrators when an alarm is raised.

Q: Are wireless networks safe?

A: No network solution is 100 percent safe, but by taking proper precautions and using strong encryption techniques, you will be able to deter the majority of attacks. If wireless security is a top priority, investing in an IDS/IPS system that is designed for wireless networks will provide the highest level of security. Wireless IPS systems can actually intercept and suppress wireless attacks in the air, preventing many of the most common attacks from reaching your access points. They can also detect rogue access points and provide wireless packet forensic analysis.

Chapter 17

E-mail Forensics

Solutions in this Chapter:

- **Where to Start?**
- **Forensic Acquisition**
- **Processing Local Mail Archives**

Introduction

E-mail or electronic mail has become a mainstay in today's society. According to the PEW Internet and American Life Project in a February-March 2007 survey, 71% of American adults use the Internet. In addition, they found 91% send or read e-mail. However, just because people use it does not mean they know how it works. Many individuals have no idea how the e-mail system works. In asking random people of varying degrees of geekdom how e-mail worked, I was shocked by some of the answers:

"I know there must be the e-postal service out there that sorts through the mail and makes sure it gets to the right places."

"E-mail is just one more way we get tracked by our bosses and forced into longer work times." (Said while typing on his BlackBerry, a common handheld e-mail device)

"I do get joke emails and then I print them and mail (snail mail) them to my family. Some of those jokes are very funny."

There was a small percentage that did know how e-mail worked, which gave me some hope I was not alone in the universe when it came to being a geek. However, this also brought some interesting thoughts to mind. Because people are generally poorly informed on how it works, they don't know how it stores data and thus what is there. Eureka we have evidence.

Where to Start?

Before you can start examining e-mail archives, you have to first understand the special language that is used when talking about e-mail. Just like the new acronyms that have become part of our everyday jargon like "lol" or "rofl", e-mail has unique words that are used to describe the smaller scale ingredients of the e-mail.

E-mail Terminology

IMAP: Internet Message Access Protocol is a method to access e-mail or bulletin board messages that are kept on a mail server making them appear and act as if they were stored locally.

MAPI: Messaging Application Program Interface is a MS Windows interface that allows you to send e-mail from inside an application. Typical applications that work with this option are word processors, spreadsheets, and graphic applications.

SMTP: Simple Mail Transfer Protocol receives outgoing mail from clients and validates source and destination addresses. It also sends and receives email to and from other SMTP servers. The standard SMTP Port is 25

HTTP: Hypertext Transfer Protocol is typically used in web mail and the message remains on the web mail server.

ESMTP: Enhanced SMTP is protocol extensions to the SMTP standard.

POP3: Post Office Protocol 3 is a standard protocol for receiving e-mail that deletes mail on the server as soon as the e-mail has been downloaded by the user. The standard port for POP3 is 110.

CC: Carbon Copy is a field in the e-mail header that directs a copy of the message to go to another recipient e-mail address.

BCC: Blind Carbon Copy is a field that is hidden from the receiver but allows for a copy of the message to be sent to the e-mail address in this field.

HELO: Communication command from client to server in SMTP e-mail delivery.

Here is an Example HELO Exchange from Wikipedia:

http://en.wikipedia.org/wiki/Simple_Mail_Transfer_Protocol

```
S: 220 www.example.com ESMTP Postfix
C: HELO mydomain.com
S: 250 Hello mydomain.com
C: MAIL FROM:<sender@mydomain.com>
S: 250 Ok
C: RCPT TO:<friend@example.com>
S: 250 Ok
C: DATA
S: 354 End data with <CR><LF>.<CR><LF>
C: Subject: test message
C: From: sender@mydomain.com
C: To: friend@example.com
C:
C: Hello,
C: This is a test.
C: Goodbye.
C: .
S: 250 Ok: queued as 12345
C: QUIT
S: 221 Bye
```

Used only if server does not respond to EHLO

EHLO: HELO Command in ESMTP clients

NNTP: Network News Transfer Protocol is used for newsgroups similar to standard e-mail. Headers are usually downloaded first in groups. The bodies are downloaded when the message is opened.

Each of these items will help you to understand the e-mail archives and become one with its evidence value. Once you understand the terminology, it is important to also understand the functions.

Functions of E-mail

E-mail, as a general rule, is designed to make communication faster between individuals. Most e-mail will allow you to do a variety of things to help you facilitate it.

- From:
- Send and receive mail
- Forward, CC, & BCC mail
- Allow attachments to be sent and received
- Save mail to disk
- Store commonly used addresses
- Sort mail into predefined folders

Each of these actions will create changes that you will have in your evidence and must be considered in the processing.

Archive Types

The next step is that you have to know what you are looking for. There are two main archive types; a local archive and a server storage archive. Most of the time, these archives will become intertwined with one another as they are not always autonomous so you will have to look for multiple tiers of the archive.

An example of this can be found with the Microsoft Exchange archives. The main archive is found in a Priv.edb file. The offline storage of the EDB file is a PST file and the offline storage for a PST file is an OST file. As you can see, these layers into your final goal of the proper evidence collection can end up becoming rather messy. Each type of archive will store data differently and e-mail makes up one of the largest types of proprietary files in the binary world.

Server Storage Archives

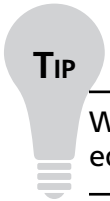
What is a server storage archive?

Server storage archives are any archive that has mixed storage for all of the clients that exist on a server. Examples of these types of archives include: MS Exchange (.EDB .STM), Lotus Notes (.NSF .ID), GroupWise (.DB), etc.

MS Exchange

When dealing with MS Exchange, it is important to remember some helpful hints.

1. Do not deal with an active Exchange server. You will want to make sure that whenever possible the Exchange server is not actively being accessed. There are many disputes to take it offline to do your image or not. One of the most successful methods is to do a backup of the server. This will maintain the best date structure for the data.
2. Always gather all the data files associated with the server. There is more than one file associated with Exchange email, so it is important to make sure you gather them all as part of your acquisition. Typically, you will find a PRIV.EDB file, PUB.EDB file, and a PRIV.STM file. These files are what create the complete archive. Although your tools might not open these files directly they will still need the reference data while they are opening the main archive. Depending on the version of Exchange you are dealing with, the files available might vary.



TIP

Watch for administrators that might change the names of the file. The Priv.edb data will be found in the larger of the two files.

3. Beware of backups and offline storage. One of the biggest headaches in dealing with server e-mail is the fact that a lot of times backups will be part of your forensic process. It is still very common for backups and archives of the enterprise mail servers to exist on tape. This can be problematic as it is a specialty to be able to process tapes. If you are not familiar with restoring tapes, it is always recommended that you go to a specialist and have them process them for you.

Lotus Notes

I have always classified Lotus Notes at a higher level archive because it is typically used in an enterprise environment. It can be easy to gather the evidence from this type of archive but difficult to extract.

1. Gather the *.NSF file
2. Gather the associated *.ID file for the archive. This is the encryption key that will allow you to open encrypted mail that is stored.

Novell GroupWise

This is not as common of a network archive as the prior mentioned archives, however, it is still found in many forensic cases. There are a couple keys to dealing with a GroupWise mail archive.

1. Do not change the structure. This may seem like an odd hint but GroupWise is not the same as the others where all the mail can typically be found in one file. It is a tad more obsessive compulsive than the other archives and it breaks its mail into post offices. This means you have to make sure the acquisition is done on the entire directory and the structure remains intact, otherwise your chances of processing through the mail located in these post offices is slim.
2. Ngwguard.db is the key file for the GroupWise structure. It is typically stored in the root of mail directory and tells GroupWise about each user account and where they are located. Other key files include gwcheck.db and wphost.db however the entire directory must be intact to do examination.

Local Level Archives

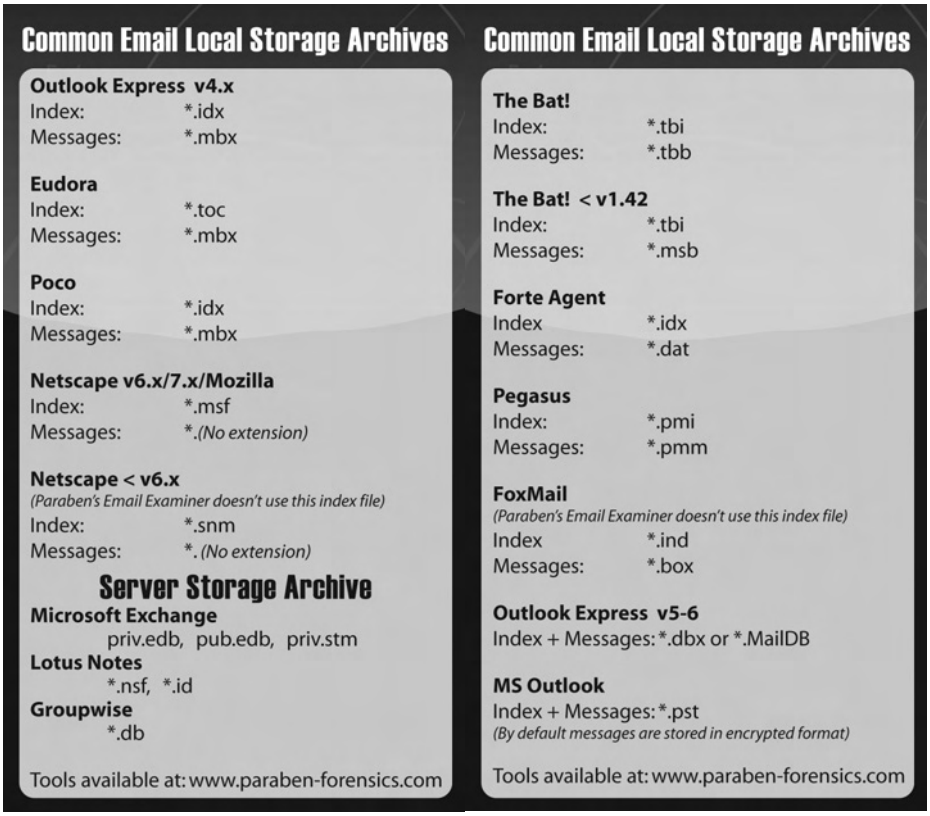
What is a local storage archive?

Local storage archives are any archives that have an archive format independent of a mail server. Examples of these types of archives include: .PST, .MBX, .DBX, etc.

The local level archives are much more diverse and can be somewhat more difficult to deal with as they are controlled more by the end user. There are still some helpful hints when dealing with local archives.

1. Always make sure you gather the entire archive. Just like with network level archives, the local archives can also be broken into multiple files that are used to store the data. Each of these files contains potential evidence and must be processed.
2. Beware the web-mail. Web-mail is very difficult to deal with as part of forensic evidence because in most cases, there is no offline archive. The data for a lot of the more popular web-mail by default is stored completely online making it difficult for a forensic examiner. If you are dealing with a web-mail archive, consult your council on the case as to the best way to approach and gain access to the servers that might contain that data.

Figure 17.1 Example of an E-mail Archive Card that is provided free from Paraben Corporation.



This card shows you the types of files that are typically associated with the e-mail archives so when you are doing seizures or examinations, you know what you should be looking for. To request cards (up to 25 per organization) e-mail emailcards@paraben.com

Ingredients of E-mail

There are some main components that will make up a mail archive. Each of these has a mutual dependence similar to if you were taking ingredients and making a cake. You could not make a proper cake without flour and eggs. With e-mail, you cannot have a proper e-mail message without a header, body, and encoding which all comes together in a single archive.

Mailbox/Archive

Message
- Header
- Body
- Encoding

Attachments

Each of these ingredients to the e-mail archive will effect your forensic examination. As part of your processing in forensics, a MD5 or other hash value will be generated as a mathematical fingerprint for the file. With e-mail archives, the problem exists that it is not just one piece or file that you are looking at but a collection of data inside.



TIP

Check your software tools prior to processing the e-mail archive on how they calculate the hash value used by the tool. The hashing mechanism should account for the e-mail header, body, and attachments and have a single hash value that represents all these components. Without looking at them as a whole, it is the same as if you were to receive a snail mail and consider only the envelope as evidence while ignoring the letter inside.

Mailbox Archive

This is the storage center or post office for the e-mail. The e-mail archive is a unique file that allows for allocated and unallocated data to live within a single logical file. In fact, during one test, I created a 1 GB PST file then proceeded to delete all messages and deleted them from the deleted items folder. The file remained 1 GB and the e-mail I deleted was all recovered by a forensic program.

Other Associated Files of the Archive

Some of the other files found in an e-mail archive include the table of contents files. These files act as a directory of the details of the mail messages. It is important to make sure when processing an e-mail archive that you process it with its associated table of contents or index file to receive the proper forensic results. Some of the common items that are stored in the table of contents or index files are:

- Main Status
- Unread
- Read

- Forwarded
- Redirected
- Flagged
- Deleted

Message

Header

The e-mail header is the envelope of the e-mail containing such information as:

- Sender E-mail Address
- Receiver E-mail Address
- Subject
- Time of Creation
- Delivery stamps
- Message Author
- CC-Carbon Copy
- BCC

All of this information can be available to you as part of your forensic analysis, but 100% of this type of data will not be found on all e-mail messages.

Body

The body is the letter of the message or the primary content.

Encoding

The encoding acts as a universal translator for the email. This is what allows different mail programs to pass data to one another even though they are not the same.

Encoding Types

MIME: Multipurpose Internet Mail Extensions is a protocol that allows non-ASCII files like video, audio, graphics, to be included in the e-mail message. In order for it to work, both sender and receiver must be able to support MIME. Most commonly used in local e-mail archive applications.

UUCODE: UNIX format for attachment encoding

BINHEX: Mac format for attachment encoding

Attachments

These are the extra items that come as supplements to the body. From pictures to files, the attachments of the e-mail archive are endless. Typically analysis of the e-mail attachments has to be done with separate tools that understand the variety of proprietary files that can be sent as attachments.

Breakdown EXAMPLE
MIME-Version: 1.0
From: Cpt Picard <cptpicard@paraben.com>
To: Beverly Crusher <docbev@hotmail.com>
Subject: Pictures of my neck in zip file
Content-Type: multipart/mixed; boundary=boundarystring --boundarystring
Content-Type: text/plain

Attached is the file neck.zip, which has been base64 encoded.
--boundarystring
Content-Type: application/octet-stream; name="neck.zip"
Content-Transfer-Encoding: base64
Content-Disposition: attachment; filename="neck.zip"

H52QLID6AJFBALJHLIHKOLNS80JOPSNLJKNLFDLSHFLSHDLFSHLKDNC809S
AOIHN3OFNSA80HLDBJSUF93HFSLBNCOISAY890EY0AHFLNC739HFOEBOASH
OFHSODIY8930…
OAIHOFIDHF8920DFNSOFNDOSGU03UQAFLASNFDLIU03WQJFOSIFH03I9AH
FDALHFNB=

--boundarystring--

Does this message have an attachment?

Answer: Yes it is a file names neck.zip

Was there anyone CC in this message?

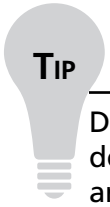
Answer: No there is no CC in the header.

This is a very simple example of an e-mail but it allows for an illustration of the basic components of the e-mail.

Forensic Acquisition

There are many tools that can process through e-mail archives. Each tool has its positive and negative points and those should be evaluated prior to purchase. However, no matter which tool you purchase, you will want to insure that you test properly and understand how it goes about its forensic validation. Since there is no standard available on how to process all the different proprietary mail formats, each tool can receive slightly different results in the processing. Here are some helpful tests for your e-mail examination tools.

1. How does it compute the hash value? Before you cross validate your tools, it will be important to find out if they are both using the same premise for validation. Some tools do not include all of the components of the mail message in the computation of the hash value. Generally the hash should include the header, body, and when applicable the attachment. It has become common for the attachment to also be extracted and hashed independently as well.
2. Was the tool designed for forensics? The processing of mail for forensics is a different process then just reading the mail archive. Your tool of choice should be able to recover deleted data from the archive as well as recover the deleted data from the archive.
3. Company support. The company should be willing to support you as a forensic examiner with good documentation explaining the process their application uses for processing as well as support for court purposes if it is required.



TIP

Deleted data is data that has been deleted from the archive's recycling bin or deleted items folder and remains in the unallocated space of the e-mail archive. Another method to create this type of data is to use the shift-delete function; this will move them directly to the unallocated space in the e-mail archive.

The following examples are processed using different tools and show what the end results should be. For complete information on any tool, please contact the vendor of the tool for that information.

Processing Local Mail Archives

The two most common e-mail archives available on local systems are Outlook Express and Outlook PST files. They are both typically found on the desktop system of the users.

Step 1-Acquisition Outlook PST file

Typically, you will do a traditional bit-stream image of the entire drive and then extract the PST file from the drive image. When extracting the PST file from the image, it is important to use multiple tools. There are many good virtual mounting programs available that allow you to mount your acquired drive and then extract a copy of the data from that drive. This is one of the better methods for extraction since some of the common methods built into the automated forensic suites will not extract a usable PST file.

Step 2-Processing

Once the file is extracted, you can select your tool for processing the proprietary e-mail archive into usable messages. I will use two tools below to illustrate the differences that can be found in processing.

Using Paraben's E-mail Examiner

Paraben's E-mail Examiner is designed to process a wide variety of e-mail archives. One of those is Outlook PST files. To process the files with this tool, a separate import engine was designed. The E-mail Examiner is shown in Figure 17.2 and the PST Converter is shown in Figure 17.3.

Figure 17.2 Paraben's E-mail Examiner

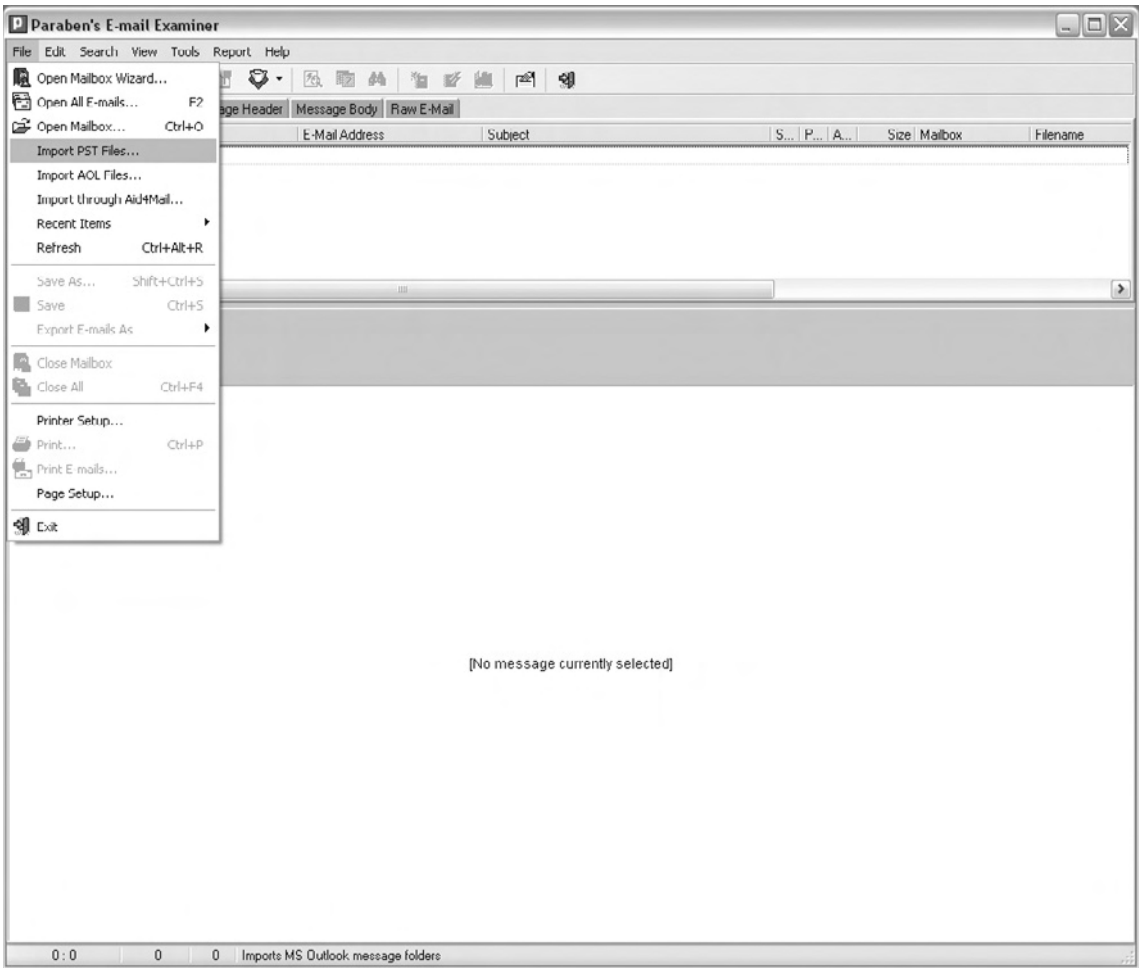


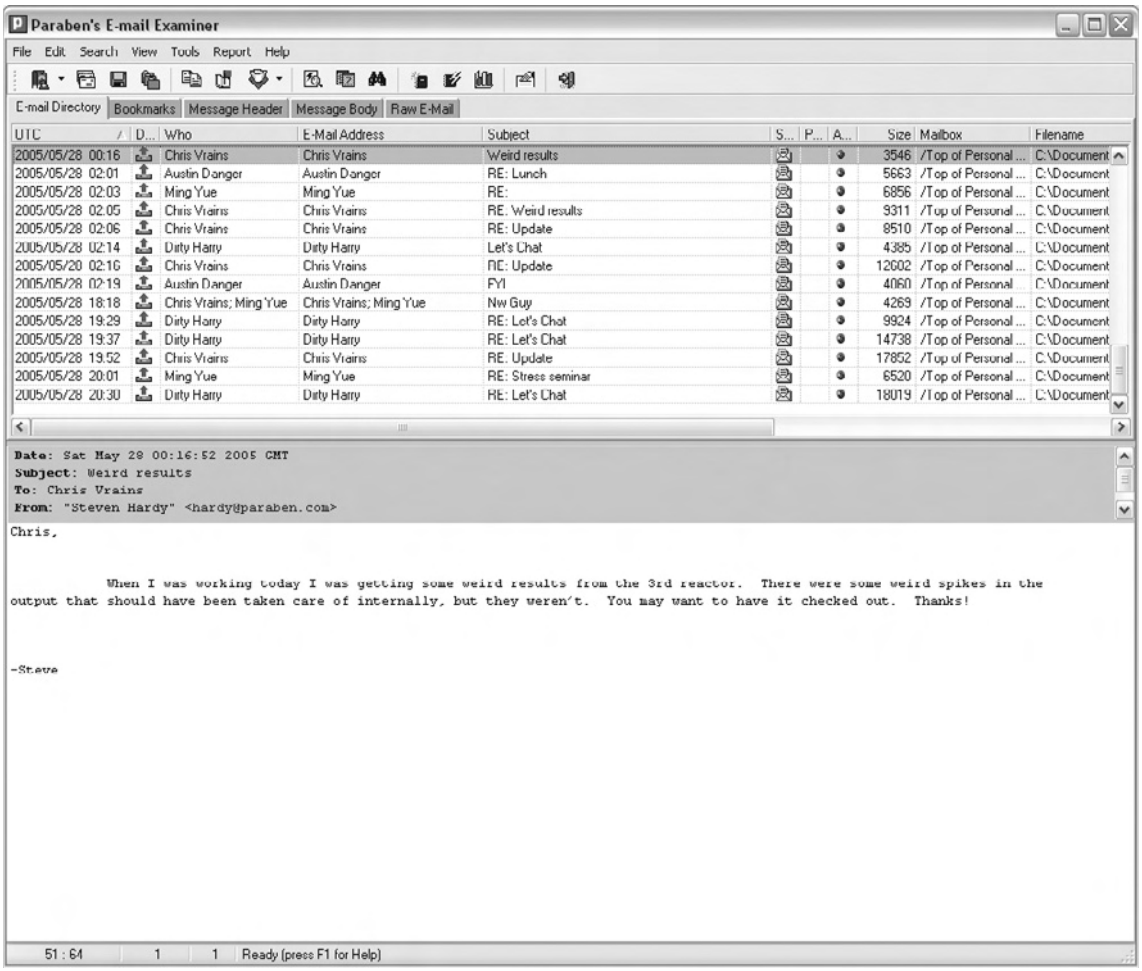
Figure 17.3 Paraben's PST Converter


PST Converter
Paraben's PST Converter reads MS Outlook message folders and exports them to generic format thet Paraben's E-mail Examiner can process. This conversion procedure doesn't affect integrity of the native MS Outlook files.
Use the "Add Files" button to manually search and select .pst files from your disk, or use "Search Disk" to list all .pst files stored on the choosen drive.
Add Files..
Search Disk...
Convert
Y:\Testing Files\EMX\PST- Special Cases from NEMX\Sent Items2.pst
Include Attachments
Save in MIME format
Recover deleted messages
Localize html image tags
Show system folder
Save to:
C:\Program Files\Paraben Corporation\E-mail Examiner\Mailb
...
Copyright © 2003 Paraben Corp.

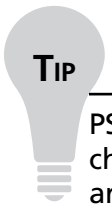
After you have selected the file to import, you are left with a variety of options for the actual processing of the archive. Each of the options listed in the screen shot will affect what you see as the ending data. The recovery of deleted messages through this engine works for both deleted and deleted-deleted data. However, once the mail archive is processed, the data that was recovered from deleted processing does not get tagged as being different from any of the other mail messages. It is important to remember this so you can look for other details that would tell you that those messages were recovered from that space such as the path.

Once the files are processed, the details will be displayed for you as seen in Figure 17.4.

Figure 17.4 Processed Files


Paraben's E-mail Examiner
File Edit Search View Tools Report Help
E-mail Directory Bookmarks Message Header Message Body Raw E-Mail
UTC D... Who E-Mail Address Subject S... P... A... Size Mailbox Filename
2005/05/28 00:16 Chris Vrains Chris Vrains Weird results 3546 /Top of Personal ... C:\Document
2005/05/28 02:01 Austin Danger Austin Danger RE: Lunch 5663 /Top of Personal ... C:\Document
2005/05/28 02:03 Ming Yue Ming Yue RE: 6856 /Top of Personal ... C:\Document
2005/05/28 02:05 Chris Vrains Chris Vrains RE: Weird results 9311 /Top of Personal ... C:\Document
2005/05/28 02:06 Chris Vrains Chris Vrains RE: Update 8510 /Top of Personal ... C:\Document
2005/05/28 02:14 Dirty Harry Dirty Harry Let's Chat 4385 /Top of Personal ... C:\Document
2005/05/28 02:16 Chris Vrains Chris Vrains RE: Update 12602 /Top of Personal ... C:\Document
2005/05/28 02:19 Austin Danger Austin Danger FYI 4060 /Top of Personal ... C:\Document
2005/05/28 18:18 Chris Vrains; Ming Yue Chris Vrains; Ming Yue Nw Guy 4269 /Top of Personal ... C:\Document
2005/05/28 19:29 Dirty Harry Dirty Harry RE: Let's Chat 9924 /Top of Personal ... C:\Document
2005/05/28 19:37 Dirty Harry Dirty Harry RE: Let's Chat 14738 /Top of Personal ... C:\Document
2005/05/28 19:52 Chris Vrains Chris Vrains RE: Update 17852 /Top of Personal ... C:\Document
2005/05/28 20:01 Ming Yue Ming Yue RE: Stress seminar 6520 /Top of Personal ... C:\Document
2005/05/28 20:30 Dirty Harry Dirty Harry RE: Let's Chat 18019 /Top of Personal ... C:\Document
Date: Sat May 28 00:16:52 2005 GMT
Subject: Weird results
To: Chris Vrains
From: "Steven Hardy" <hardy@paraben.com>
Chris,
When I was working today I was getting some weird results from the 3rd reactor. There were some weird spikes in the output that should have been taken care of internally, but they weren't. You may want to have it checked out. Thanks!
-Steve
51 : 64 1 1 Ready (press F1 for Help)

There are many things that you will notice once the archive is processed from deleted data recovery to messages with attachments.



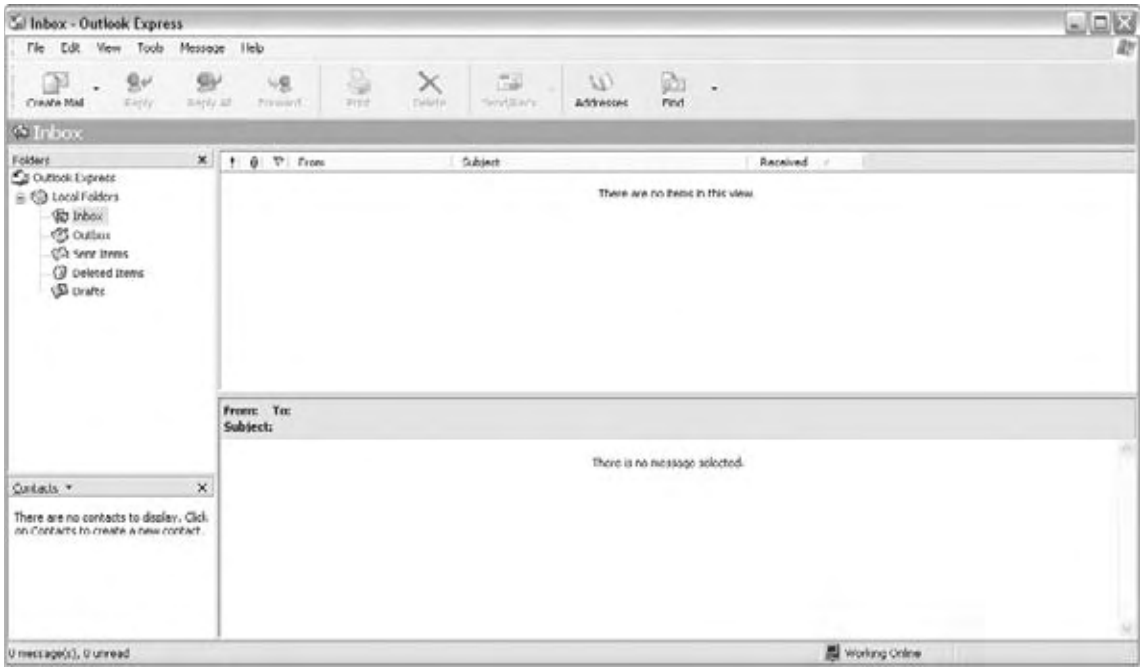
TIP

PST files that are made in MS Outlook 2003 or greater have had a format change occur with their data structure. The previous versions had a smaller archive size limit and typically would process through easier than the new archives. Pay attention to what version of MS Outlook your suspect was using so you can look for offline archives (OST files) or look for other PST files besides their main archive.

Using MS Outlook for processing Outlook Express files

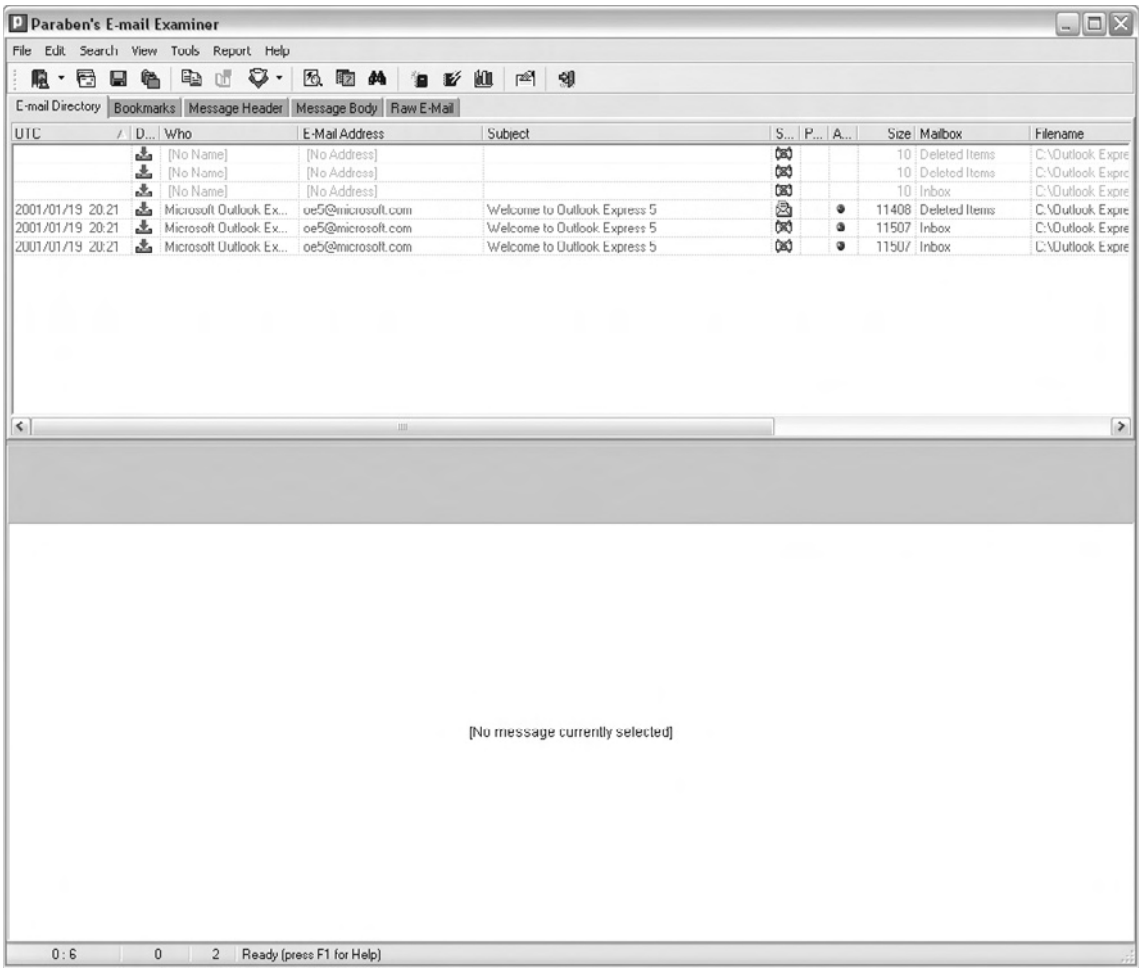
Some people prefer to use tools that are the mail clients for processing the data associated with the archive. This can be problematic because these tools are not specifically designed for forensics. So, much of the deleted data would be missed in the processing. To illustrate this I have processed the same archive (Outlook Express) with both the mail application (Figure 17.5) and a forensic application (Figure 17.6).

Figure 17.5 Mail Application



According to the mail application there is no data in the e-mail archive.

Figure 17.6 Processing with Forensic Tool


Paraben's E-mail Examiner
File Edit Search View Tools Report Help
E-mail Directory
Bookmarks
Message Header
Message Body
Raw E-Mail
UTC
D...
Who
E-Mail Address
Subject
S...
P...
A...
Size
Mailbox
Filename
[No Name] [No Address] 10 Deleted Items C:\Outlook Expre
[No Name] [No Address] 10 Deleted Items C:\Outlook Expre
[No Name] [No Address] 10 Inbox C:\Outlook Expre
2001/01/19 20:21 Microsoft Outlook Ex... oe5@microsoft.com Welcome to Outlook Express 5 11408 Deleted Items C:\Outlook Expre
2001/01/19 20:21 Microsoft Outlook Ex... oe5@microsoft.com Welcome to Outlook Express 5 11507 Inbox C:\Outlook Expre
2001/01/19 20:21 Microsoft Outlook Ex... oe5@microsoft.com Welcome to Outlook Express 5 11507 Inbox C:\Outlook Expre
[No message currently selected]
0 : 6
0
2
Ready (press F1 for Help)

Once the forensic tool has processed the archive, a variety of messages were recovered.

Processing Server Level Archives.

As discussed previously, there are many different files to look for when processing a server level archive. Depending on which mail server was used, you will need to gather different data as previously discussed.

Step 1 Acquisition

The acquisition stage for a server archive is different than with the smaller local stores as you do not typically do a bit-stream image and then extract the archive. Instead, in most cases you can just acquire the appropriate files where the archive data is stored. Although this is not a traditional forensic method, it is very common based on the structure of the network archive and size.

Step 2 Processing

There is not a wide range of tools available for network level archives. Most tools are not design specifically for forensic processing, so you are limited in your choices of tools if you want to stay just in forensic software. However, you do have other tool options available that are designed for restoring archives for review.



WARNING

It is very common for MS Exchange archives to be corrupted. Causes can include backup issues, size of the archive or any number of things. For most corruption, the common method is to use EseUtil.exe that comes with MS Exchange. Although this tool may work in some cases it is not recommended as a first resort. Check your forensic tools for corruption repair capabilities and try using them first. This will allow them to have a pure access to the files as EseUtil.exe is known to change data.

Using OnTrack PowerControls

"Ontrack PowerControls is a simple, yet powerful software tool for copying, searching, recovering and analyzing email and other mailbox items directly from Microsoft Exchange server backups, un-mounted databases (EDB) and Information Store files." **http://www.ontrackpowercontrols.com**

PowerControls is one of the better tools available for processing MS Exchange files. It recovers both active and deleted data and can work on a variety of versions of MS Exchange. Figures 17.7, 17.8, and 17.9 show the data that has been processed.

Figure 17.7 Process MS Exchange PRIV.EDB file.

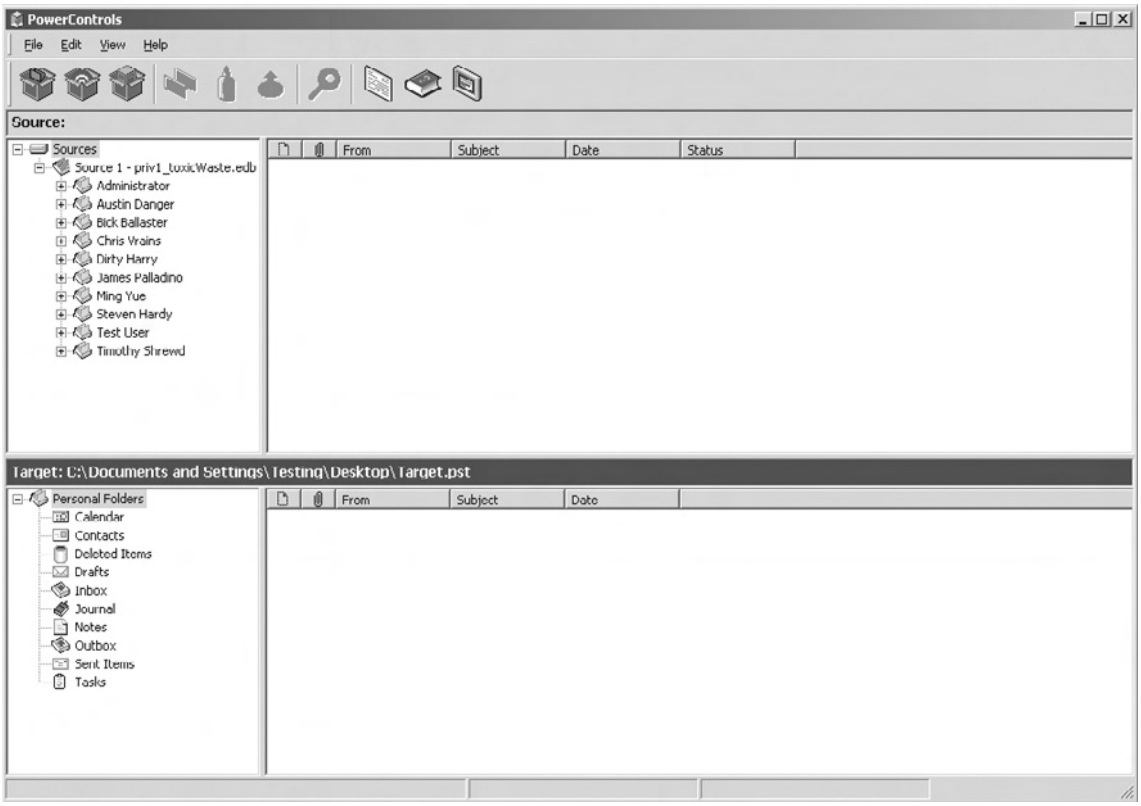


Figure 17.8 The display of an individual account in the MS Exchange file.

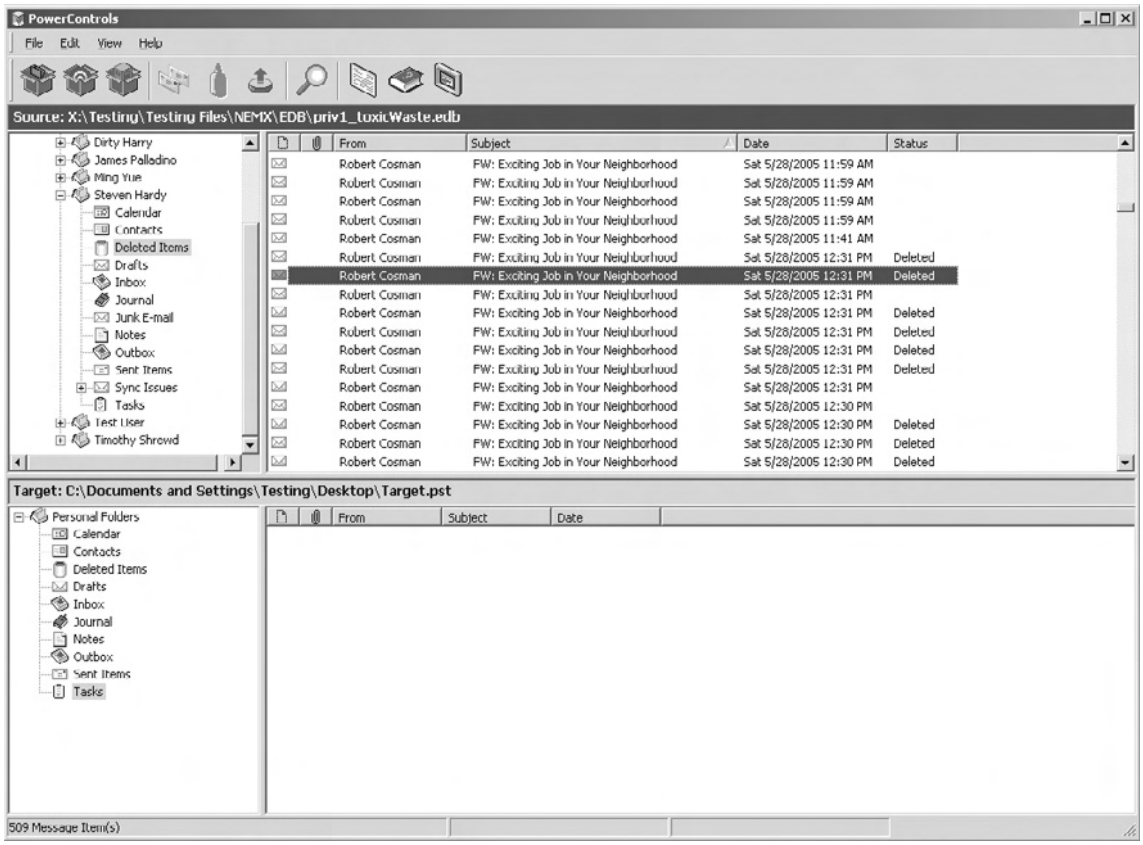
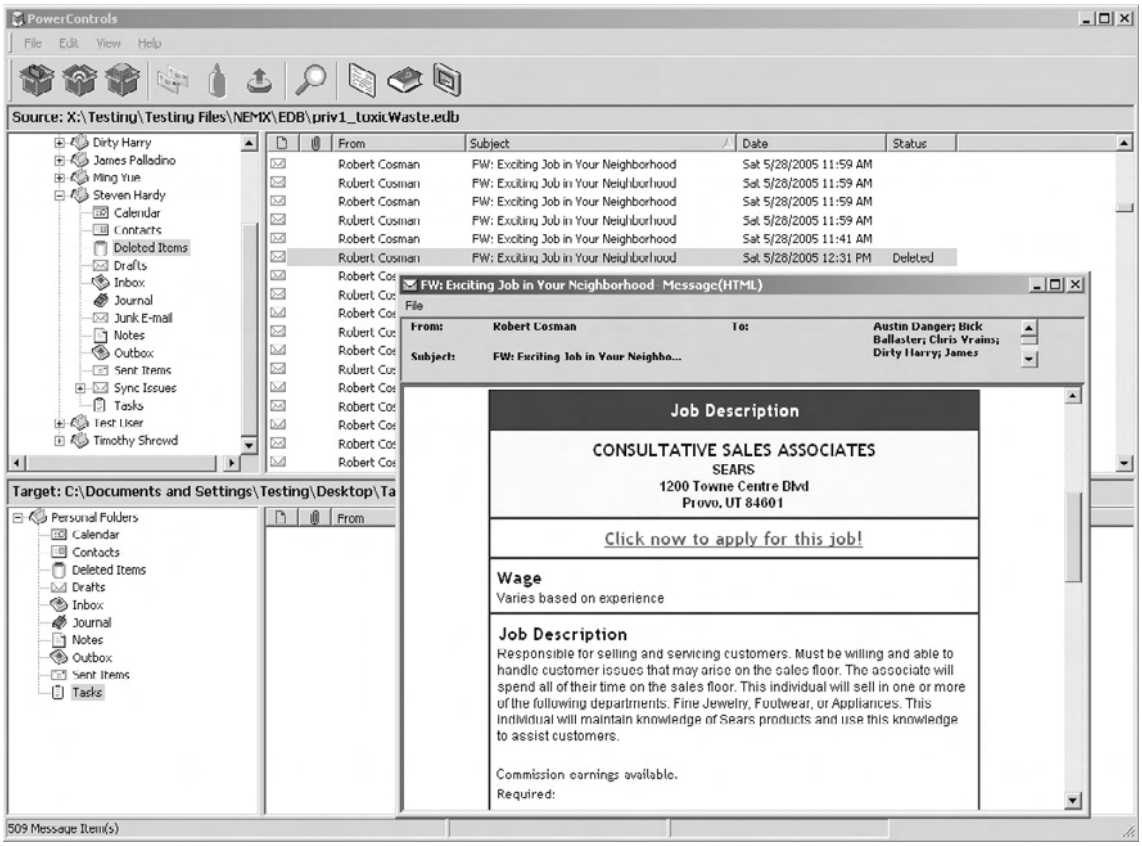


Figure 17.9 Convenient message viewer for the review of the content data in the e-mail archive.


PowerControls
File Edit View Help
Source: X:\Testing\Testing Files\NEMX\EDB\priv1_ToxicWaste.edb
Dirty Harry
James Palladino
Ming Yue
Steven Hardy
Calendar
Contacts
Deleted Items
Drafts
Inbox
Journal
Junk E-mail
Notes
Outbox
Sent Items
Sync Issues
Tasks
Test User
Timothy Shrewd
From
Subject
Date
Status
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 11:59 AM
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 11:59 AM
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 11:59 AM
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 11:59 AM
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 11:41 AM
Robert Cosman FW: Exciting Job in Your Neighborhood Sat 5/28/2005 12:31 PM Deleted
FW: Exciting Job in Your Neighborhood Message(HTML)
File
From: Robert Cosman
To: Austin Danger; Bick Ballaster; Chris Vrains; Dirty Harry; James
Subject: FW: Exciting Job in Your Neighbo...
Job Description
CONSULTATIVE SALES ASSOCIATES
SEARS
1200 Towne Centre Blvd
Provo, UT 84601
Click now to apply for this job!
Wage
Varies based on experience
Job Description
Responsible for selling and servicing customers. Must be willing and able to handle customer issues that may arise on the sales floor. The associate will spend all of their time on the sales floor. This individual will sell in one or more of the following departments: Fine Jewelry, Footwear, or Appliances. This individual will maintain knowledge of Sears products and use this knowledge to assist customers.
Commission earnings available.
Required:
Target: C:\Documents and Settings\Testing\Desktop\Ta
Personal Folders
Calendar
Contacts
Deleted Items
Drafts
Inbox
Journal
Notes
Outbox
Sent Items
Tasks
509 Message Item(s)

Using Paraben's Network E-mail Examiner (NEMX)

NEMX is also a tool that can be used to process MS Exchange archives as well as Lotus Notes, and GroupWise. Built into the tool is a corruption repair utility that will also save some time in processing by attempting to bypass corruption and moving on to read the rest of the archive allowing to keep the data in its original state. Figures 17.10, 17.11, and 17.12 show some examples from processing a MS Exchange PRIV.EDB file.

Figure 17.10 A fully processed MS Exchange file including server level information.

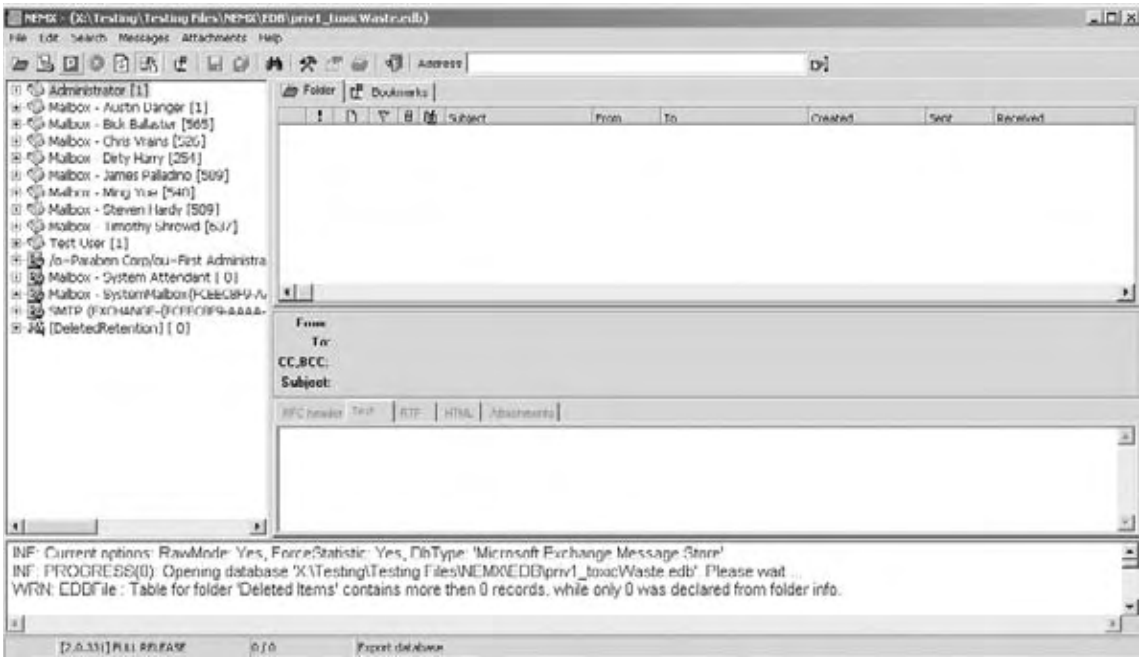
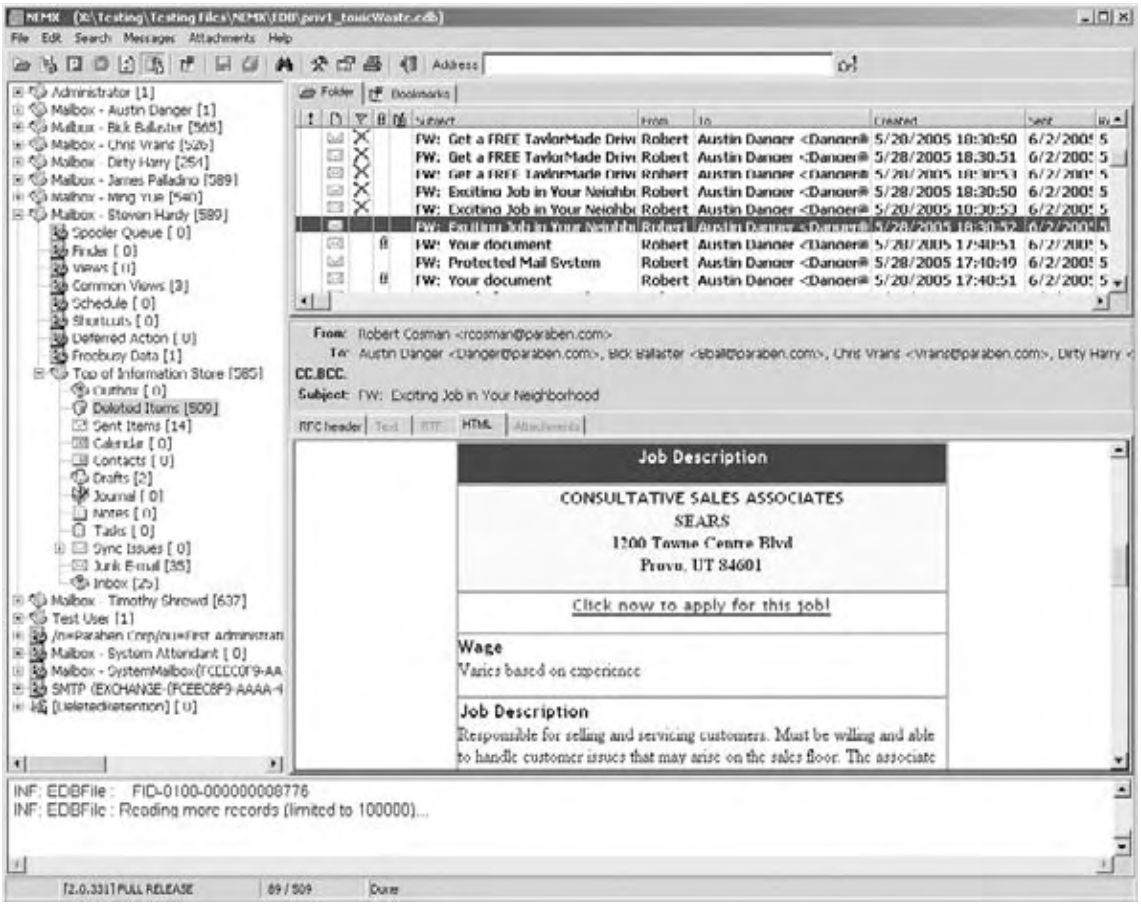

Administrator [1]
Mailbox - Austin Danger [1]
Mailbox - Bick Ballaster [565]
Mailbox - Chris Vrains [526]
Mailbox - Dirty Harry [254]
Mailbox - James Palladino [589]
Mailbox - Ming Yue [540]
Mailbox - Steven Hardy [589]
Mailbox - Timothy Shrewd [637]
Test User [1]
/o=Paraben Corp/ou=First Administra
Mailbox - System Attendant [0]
[DeletedRetention] [0]
From
To
CC,BCC
Subject
INF: Current options: RawMode: Yes, ForceStatistic: Yes, DbType: 'Microsoft Exchange Message Store'
INF: PROGRESS(0): Opening database 'X:\Testing\Testing Files\NEMX\EDB\priv1_toscWaste.edb' Please wait...
WRN: EDBFile : Table for folder 'Deleted Items' contains more then 0 records, while only 0 was declared from folder info.

Figure 17.11 Tree view of the data typically associated with an MS Exchange Priv.edb file.


Administrator [1]
Mailbox - Austin Danger [1]
Mailbox - Bick Ballaster [565]
Mailbox - Chris Vrains [526]
Mailbox - Dirty Harry [254]
Mailbox - James Palladino [589]
Mailbox - Ming Yue [540]
Mailbox - Steven Hardy [589]
Mailbox - Timothy Shrewd [637]
Test User [1]
/o=Paraben Corp/ou=First Administrat
Mailbox - System Attendant [0]
Mailbox - SystemMailbox{FCEEC8F9-AA
SMTP (EXCHANGE-{FCEEC8F9-AAAA-4
[DeletedRetention] [0]

Figure 17.12 Opening of the data associated with the user account in the MS Exchange file.


File Edit Search Messages Attachments Help
Administrator [1]
Mailbox - Austin Danger [1]
Mailbox - Bill Ballaster [565]
Mailbox - Chris Vrains [526]
Mailbox - Dirty Harry [264]
Mailbox - James Palladino [589]
Mailbox - Ming Yue [540]
Mailbox - Steven Hardy [589]
Spooler Queue [0]
Finder [0]
Views [11]
Common Views [3]
Schedule [0]
Shortcuts [0]
Deferred Action [0]
Freebusy Data [1]
Top of Information Store [585]
Outbox [0]
Deleted Items [509]
Sent Items [14]
Calendar [0]
Contacts [0]
Drafts [2]
Journal [0]
Notes [0]
Tasks [0]
Sync Issues [0]
Junk E-mail [35]
Inbox [26]
Mailbox - Timothy Shrewd [637]
Test User [1]
Mailbox - System Attendant [0]
SMTP (EXCHANGE-(FCEEC8F9-AAAA-
FW: Get a FREE TaylorMade Driv Robert Austin Danger <Danger@ 5/20/2005 10:30:50 6/2/2005
FW: Get a FREE TaylorMade Driv Robert Austin Danger <Danger@ 5/28/2005 18:30:51 6/2/2005
FW: Exciting Job in Your Neighbo Robert Austin Danger <Danger@ 5/28/2005 18:30:50 6/2/2005
FW: Exciting Job in Your Neighbo Robert Austin Danger <Danger@ 5/20/2005 10:30:53 6/2/2005
FW: Your document Robert Austin Danger <Danger@ 5/20/2005 17:40:51 6/2/2005
FW: Protected Mail System Robert Austin Danger <Danger@ 5/28/2005 17:40:49 6/2/2005
FW: Your document Robert Austin Danger <Danger@ 5/20/2005 17:40:51 6/2/2005
From: Robert Cosman <rcosman@paraben.com>
CC,BCC,
Subject: FW: Exciting Job in Your Neighborhood
RFC header
HTML
Job Description
CONSULTATIVE SALES ASSOCIATES
SEARS
1200 Towne Centre Blvd
Provo, UT 84601
Click now to apply for this job!
Wage
Varies based on experience
Job Description
Responsible for selling and servicing customers. Must be willing and able
to handle customer issues that may arise on the sales floor. The associate
INF: EDBFile : FID-0100-000000008776
INF: EDBFile : Reading more records (limited to 100000)...
72.0.3317 FULL RELEASE
69 / 509
Done

TIP

Many examiners expect to find the same details in the Priv.edb file as they do the offline store *.PST file. This is not always the case. When trying to do a cross validation, a message to message comparison should be done by viewing the area in the MS Exchange archive for that user called Top of Information Store. This is the closest direct comparison you will find to the *.PST that may exist as offline backup.

Deleted E-mail Recovery

The recovery of deleted e-mail messages can vary greatly on the e-mail client that was being used. Typically, you will find a couple weeks of deleted data that can still be recovered from an archive. Here are a few examples of how deleted e-mail works.

Eudora Mail

In Eudora, messages for deletion are tagged for deletion and are no longer visible in the mailbox. These messages, however, are still in the "trash" folder and remain there until explicitly instructed to empty the trash folder.

Outlook PST

In Outlook, data is taken from the active part of the archive to a recycling bin. From that point, the recycling bin is emptied based on the user's preferences. Once it is emptied, it will go into the unallocated space of the mail archive. Here it can sit for a period of weeks. Depending on the size of the archive the recovery of this data will vary greatly.

Network Archives

Depending on the network level archive, you will have a variety of results on deleted mail recovery. It is common to recover a good percentage of deleted e-mail.

Overall, the processing of e-mail can be made simple by following guidelines and having the proper expectations.

"I know my e-mail goes through my computer, but beyond that I don't know and I don't care as long as it goes. How would I know? My clock on my VCR still blinks." (Survey Participant)

Now that you know more than the average bear when it comes to e-mail, you are well on your way to becoming an e-mail forensic examiner. Always remember to keep learning, as this type of information changes constantly.

Chapter 18

Steganography and Application Password Crackers

Solutions in this chapter:

- Classification of Steganography
- Six Categories of Steganography in Forensics
- Types of Steganography
- Application of Steganography
- Hiding in Network Packets
- Issues in Information Hiding
- Steg Tools
- Steganography vs. Watermarking
- Attacking Watermarking
- Detecting and Attacking Steganography
- Application Password Cracking

☑ Summary

Introduction

In this chapter we will discuss the technique of steganography and application password cracking. First we will discuss the concept of steganography. Steganography is defined as "hiding information in plain site." Steganography differs from cryptography in the fact that steganography does not take a "plain" text form as input and scramble it into a "cipher" text like cryptography does. Furthermore, steganography does not alter the form of the information it is hiding. Quite simply, you use steganography to place information inside a file.

There are two types of files that are used in steganography: the file that you use to hold information that can be in the form of text or another file called the carrier file; and the file that is to be inserted into the carrier, known as the hidden file. The carrier file is also referred to as the host file in some steganography documentation. Now that we have explored what steganography is, let's take a look at some of the past history of steganography.

History of Steganography

Steganography comes from the joining of two words. These words are of Greek origin: *steganos*, meaning covert or secretive; and *graphie*, meaning writing. Putting the two words together means "*covered writing*."

The Greeks

In *The Histories*, Herodotus documented several instances where steganography was used. One such case involved a Greek named Histaiaeus, who wanted to encourage Aristagoras of Miletus to revolt against the Persian king, and managed to do so in a rather inventive way. In order to pass these instructions securely, Histaiaeus shaved the head of his messenger, wrote the message on his bare scalp, and then waited for the hair to grow back. When the messenger reached his destination—and Aristagoras—his head was shaved and the message read. Although this certainly isn't the quickest method of communication, it was very effective since the messenger was able to pass guard inspections without harassment since he was carrying nothing suspicious.

The Chinese

The Chinese used steganography during the Yuan dynasty to revolt when China was ruled by the Mongolian people. The leaders of the time, unhappy at having to submit to foreign rule, created an inventive method of coordinating the rebellion without being discovered. The rebel leaders, knowing that the Moon Festival was drawing near, ordered the making of special cakes called "moon cakes." Baked into each moon cake was a message with the outline of the attack. On the night of the Moon Festival, the cakes were dispersed,

the messages read, and the rebels successfully attacked and overthrew the government. What followed was the establishment of the Ming dynasty. Today, moon cakes are eaten to commemorate this legend.

The Culpers

Steganography played a role in the Revolutionary War and helped George Washington on many occasions. The code ring consisted of five people and used the codename Samuel Culper. Several trial and error attempts at secret communication prompted the Culpers to adopt invisible ink as one of their security precautions. The invisible ink was used on a blank piece of paper, and after the message was written, was reinserted into a ream of new, blank paper. Washington knew that by counting from the top down to a specific sheet how to find the hidden message. He would then apply a second solution to make the ink appear. A concern of Washington's was that carrying around blank sheets of paper would create suspicion, so he ordered that the invisible ink also be used on regular messages, and that the secret message be written between the lines or under the message.

Civil War Rugs

A form of steganography was used prior to the Civil War to aid slaves in escaping to freedom. In the 1800s, the underground railroad was one of the main escape routes used by slaves. Quilts, which were hung outside to dry, became an ideal way to display information inconspicuously. The quilts would have special patterns sewn into them that would convey messages to prepare or provide direction to escaping slaves who knew what to look for.

World War I

During World War I there were several periods and instances where steganography was used with success. One method was called a Turning Grille; this took Cardano's Grille method and enhanced it. The Turning Grille looked like a normal grille (a square sheet of cardboard) divided into cells, with some of the cells punched out. To use the Turning Grille, the encoder would write the first sequence of letters, and then rotate the grille 90 degrees, write the second sequence of letters, and so on, rotating the grille after each sequence.

There was also a record during the war of a cable censor who received the message "Father is dead." The censor changed the message to "Father is deceased." A reply came back: "Is Father dead or deceased?"

World War II

Steganography, and its prevention, were also prevalent in World War II. After the attack on Pearl Harbor the United States enacted a censorship organization. This organization worked to think of ways that coded messages could be passed in the open and took steps to stop them or, if there was a message, destroy the code.

Chess games were banned by mail, crossword puzzles were examined or removed from correspondence and newspaper clippings, and student's grades were removed.

Stamps were removed and replaced with ones of equal value but different denominations or numbers. Children's pictures were replaced, Xs and Os from love letters were removed, and of course blank paper was replaced and tested for invisible ink. Mass media was censored, too. Telephone and telegraph requests for special songs were not allowed and mail-in song requests would be held for a random amount of time before being played.

The Vietnam War

Another, and well-publicized, example of steganography happened during the height of the Vietnam War. Commander Jeremiah Denton, a naval pilot who had been shot down and captured, was paraded around by the North Vietnamese in front of the news as part of a propaganda event. Knowing that he was under the proverbial microscope and unable to say anything openly critical about his captors, he blinked his eyes in Morse code, spelling out T-O-R-T-U-R-E, as he spoke to the media.

Terrorists

It is speculated that the terrorists who made up the "cell" located in the United States used steganography to communicate information between the leaders of Al-Qaeda and themselves. Part of this speculation is from captured computer files that have been found during various raids. It is really impossible to know if the "cell" did actually use steganography to communicate, but equally difficult is proving they did not.

The Future of Steganography

One powerful concept of steganography we will discuss in this chapter is the technique of hiding messages or information within a binary file. This is a powerful way to hide information, and short of reverse engineering, or disassembling the code there really is no way to tell if this technique is being used. Fortunately, the technique largely exists in the minds of a select few, and we have no indications that it is being used widespread. Refer to the sidebar, "Notes from the Underground," for more.

Classification of Steganography

There are essentially three classifications of Steganography today. The first one is the technique of inserting information inside another file. The second is the process that takes the composition of the file, and substitutes information. The third is one that uses a different model, because it creates a file with the intent of hiding information; an example of this is the capabilities you can use at the Web site www.spammimic.com.

Background Information to Image Steganography

The concept of image steganography is when a file that is some form of image like a bmp, jpg, tif, or gif file is used to hide information. The information that is hidden can be in the form of a file or of an actual message. The process takes the user supplied message or image, and hides it in the file. There are several techniques for doing this that we will look at next. The main techniques are insertion, substitution, and creation. As we will see next, each has its own advantages and disadvantages when hiding information.

Insertion

In this method of steganography, the objective of the hiding technique is to place the information to be hidden inside the carrier or host file. This is done without any alteration of the existing information of the file. The problem with this is that when the information is inserted it increases the size of the original file, and makes it easy to detect the presence of the steg. You could use this technique if you made sure to wipe the original file once the steg file is created. As long as you did not leave a trace (hence, wipe!) of the original file on the machine it would be very difficult to determine the file had steg, because without the original file, traditional integrity checks like hashing would not have a baseline with which to compare.

Substitution

You might see some steganography references refer to this as transformation-based, but for this text we will refer to it as substitution. A popular way to hide information with this technique is to use the least significant bit (LSB) and noise manipulation method. This technique works by taking the binary data of the host, or carrier file, and changes the LSB of each pixel. The right-most portion of the number is the LSB, and since it is the LSB, the manipulation of it has little effect on the original number or the original image; therefore, we can modify all the LSBs in the file with little or no change in the original file. The process works by taking the file that is to be hidden, and breaking it into chunks of binary data. Once the file is broken down, the resulting chunks of binary data are written to and stored in the LSBs throughout the file. The following example shows an 8-digit binary number being encoded using LSB.

Example 1

Pixel	00	11	00	11
Insert	0010			
Pixel	00	1**0**	0**1**	10

As you can imagine, working with 8-bit colors can be a challenge; that is why images that use 24-bit color are the ones that are preferred for the best results. With a 24-bit image,

you can get usage of 3 LSBs per pixel. The reason the insertion technique works so well is the human eye can distinguish only six bits of color, and since the technique is working on the LSB of an 8-bit number. The changes are virtually undetectable.

Creation

A cover file is *generated* for the sole purpose of concealing a secret message. In this technique a picture is created of something innocent that can be passed to another person. The innocent picture is the cover that provides the mechanism for conveying the message.

Another, more modern form of generation of new files is the program Spam Mimic. Spam Mimic embeds a text message within a rather daunting piece of spam that can be e-mailed to an intended recipient. Although this generated spam doesn't make a whole lot of sense, it makes enough sense to be believable, and that's all that's needed. An excellent reference for this is the book *Disappearing Cryptography: Being and Nothingness on the Net*, by Peter Wayne (AP Professional, San Diego, 1996). The book also includes code to create this type of steganography.

Six Categories of Steganography in Forensics

In all methods of steganography, something is being done to conceal a message. Naturally, these actions or techniques can be separated and analyzed to learn what's happening during the whole process. The six categories of steganography are substitution system techniques, transform domain techniques, spread spectrum techniques, statistical method techniques, distortion techniques, and cover generation techniques. We'll look at each of these, one at a time.

Substitution System

Substitution system steganography replaces redundant or unneeded bits of a cover with the bits from the secret message. Several available steganography tools use the Least Significant Bit (LSB) method of encoding the secret message. LSB works like this. In a digital cover file, there is a tremendous amount of wasted or redundant space. It is this space that the steganography program will take advantage of and use to hide another message (on the bit level) within the digital cover.

Transform Domain Techniques

The transform domain technique is also very effective, but is a little trickier to explain. Basically, transform domain techniques hide message data in the "transform space" of a signal. If you're saying to yourself, "Huh?", hold on and I'll explain. Every day on the Internet, people send pictures back and forth, and most often they use a JPEG format. JPEGs are interesting in that they compress themselves when they close. In order for this to take place, they have to get rid of excess data (excess bits) that would otherwise prevent them from

compressing. During compression, a JPEG will make an approximation of itself to become smaller. This change transforms space; thus, it can be used to hide information.

Spread Spectrum Techniques

In this section we briefly define two types of spread spectrum techniques: direct sequence and frequency hopping.

In direct sequence spread spectrum, the stream of information to be transmitted is divided into small pieces, each of which is allocated to a frequency channel spread across the spectrum. A data signal at the point of transmission is combined with a higher data-rate bit sequence (also known as a chipping code) that divides the data according to a spreading ratio. The redundant chipping code helps the signal resist interference and also enables the original data to be recovered if data bits are damaged during transmission.

Frequency hopping is when a broad slice of the bandwidth spectrum is divided into many possible broadcast frequencies. In general, frequency-hopping devices use less power and are cheaper, but the performance of direct sequence spread spectrum systems is usually better and more reliable.

Statistical Methods

Statistical methods use what's called a "1-bit" steganographic scheme. This scheme embeds one bit of information only in a digital carrier and thus creates a statistical change, even if it's only a slight one. A statistical change in the cover indicates a 1, whereas a cover left unchanged indicates a 0. This system works based on the receiver's ability to distinguish between modified and unmodified covers.

Distortion Techniques

The distortion method of steganography creates a change in a cover object to hide information. The secret message is recovered when the algorithm compares the changed and/or distorted cover with the original.

Cover Generation Methods

Cover generation methods are probably the most unique of the six types. Typically, a cover object is chosen to hide a message in, but that's not the case here. A cover generation method actually creates a cover for the sole purpose of hiding information. Spam Mimic is an excellent example of a cover generation method.

Types of Steganography

Now that we've looked at the categories of stego, let's look at the two types of stego, linguistic and technical.

Linguistic Steganography

Linguistic steganography can be described quite simply as any form of steganography that uses language in the cover. A number of forms of linguistic steganography are covered in the next sections, but the two most basic categories are open codes and text semagrams.

A program that takes advantage of linguistic steganography is NICETEXT, which uses the technique of linguistic steganography in a very inventive way. The goal of NICETEXT is to provide a program that can transform cipher text (encrypted text) into text that looks like natural language, while still providing a cover for the original cipher text. An added benefit of this type of program is that it can be applied to many different languages.

Text Semagrams

Text Semagrams work with graphical modifications of the text, such as type-spacing and offsetting. They concern details that are tiny, but nonetheless visible. Certain methods work without text as well. These are called real semagrams.

Technical Steganography

Technical steganography is a little broader in scope since it doesn't necessarily deal with the written word, even though it communicates information. Technical steganography is the method of steganography where a tool, device, or method is used to conceal the message. Let's have a look at some technical steganography methods.

Embedding Methods

The technical challenges of data hiding are formidable. Any "holes" to fill with data in a host file, either statistical or perceptual, are likely targets for removal by lossy compression. The key to successful data hiding is the finding of holes that are not suitable for exploitation by compression algorithms.

Least Significant Bit

Least Significant Bit (LSB) is the substitution method of steganography where the right-most bit in a binary notation, the bit with the least impact on the binary data, is replaced with a bit from the embedded message. This method provides "security through obscurity," a technique that can be rendered useless if an attacker knows the technique is being used.

Transform Techniques

Three main types of transform techniques can be used when embedding a message in steganography. They are the Discrete Cosine Transform (DCT), the Discrete Fourier Transform, and the Wavelet Transform, which simply embeds the secret data in the cover during the *transformation* process in various ways.

Spread-Spectrum Encoding

Spread-spectrum encoding is the method of hiding a small or narrow-band signal (a message) in a large or wide band cover. The foundation of this process begins with a spread spectrum encoder. The encoder works by modulating a narrow band signal over a carrier. The carrier's signal is continually shifted using a noise generator and a secret key that makes the noise seem random. The message is embedded in the existing noise of the carrier signal. This spreads the narrow signal over a wide area, decreasing the density of the hidden signal and making it much more difficult to detect within the overall carrier signal.

Perceptual Masking

This form of steganography occurs when one signal, or sound, becomes imperceptible to the observer because of the presence of another signal. This method also exploits the weaknesses of the human visual and auditory systems. A common example that almost everyone has seen is in spy films where someone is trying to communicate with someone else in a room they know is bugged. Usually the secret agent will turn up the stereo, run the shower, or do some other innocent (but noisy) task that allows a whispered conversation to take place without being heard.

Notes from the Underground…

Raising the Bar in Information Hiding

At the Blackhat USA conference in 2004 in Las Vegas, two researchers, Rakfan El Khalil and Angelos D. Keromytis, from Columbia University in New York gave a presentation based on their research, and a whitepaper that they had published. They presented an advanced concept in steganography that to date had little research published on it. Part of their presentation was the demonstration of a tool that they called "Hydan," which is an old English word that means to hide or conceal. The focus of this tool is to hide information inside of binary files. It accomplishes this by taking a binary file and analyzing the machine code of the file. Once the tool finishes the analysis, the tool will take a message, and attempt to hide it in the machine code of the file. The tool deploys a registry swapping, and instruction substitution methodology. In a registry swapping methodology, the tool will take an instruction like *mov% esi,% esi* and change it to *mov %eax, %eax,* and in an instruction substitution, the tool will take *and %esi, %esi* and change it to *or %eax, %eax*.

Continued

The last update to the Hydan tool was announced in 2004, so this powerful technique does not seem to be available today, at least not to the general public. This is a good thing, because the power of a tool like this and the implications for the governments of the world would be significant, not to mention the Department of Defense (DoD) and other organizations that rely on protecting information. There really is not a sufficient way to scan through binary types of files, and look for hidden information. So, for the time being this powerful technique seems to be a proof of concept tool, and that for most is a good thing.

Application of Steganography

When you get right down to it, digital steganography is really just the digital manipulation of 1s and 0s. This fact allows for several choices of cover media since we know that whether it be a document, picture, video, or even a file system, in the end it's all just binary. In this next section, we'll look at how stego manipulation can be applied to a wide variety of digital media.

Still Images: Pictures

The methods of steganography are quite varied. With still images, LSB insertion and spread spectrum techniques commonly are used. Texture block is another method. Texture block uses low bit-rate data hiding and is accomplished by copying a region from a random texture pattern in a picture to an area of similar texture resulting in a pair of identically "textured" regions in a picture.

Moving Images: Video

Steganography, when applied to a video file such as an .avi or .mpg, typically uses Discrete Cosine Transform (DCT) manipulation, meaning that as the video is compressed or decompressed, secret data can be added during the transformation process.

Audio Files

The following are various audio file methods:

- **LSB insertion**. The bit with the least impact on the binary data is replaced with a bit from the embedded message.
- **Differential phase variation**. The sound file is divided into blocks, and the initial phase of the sound file is modified with the secret message, preserving the following relative phase shifts. This is an effective method given it has a low signal-to-noise ratio.

- **Spread spectrum schemes**. The method of hiding a small- or narrow-band signal, the secret message, in a large- or wide-band cover.
- **Adding echo to the audio signal**. A slight echo is added to the signal using two different delays to encode the 1 and 0 bits. The echo is too slight to be perceived by the human ear.

Text Files

Three methods exist for text files:

- **Open-space**. The use of white space, between words or sentences, to hide data.
- **Syntactic**. Uses modification of the word order or punctuation to hide a message.
- **Semantic**. Uses synonyms to encode a secret message. (See the earlier example, "Is Father dead or deceased?")

Steganographic File Systems

This is another particularly interesting way of applying steganography. Let's take a closer look. A steganographic file system is a method of storing files in such a way that it encrypts data and hides it so well that it can't be proven it's there.

A steganographic file system can

- Hide users' documents in other seemingly random files.
- Allow the owner to give names and passwords for some files while keeping others secret.
- Behave like a second layer of secrecy. Encrypted files are out in the open and visible, but not understandable. Stego files aren't even visible, and an outsider can't look for files that "aren't there."

Hiding in Disk Space

Three main methods exist for hiding data in disk space: unused sectors, hidden partitions, and slack space. In this section, we'll take a closer look at how each is carried out.

Unused Sectors

Similar to the method used in the section, "Steganographic File Systems," tools that hide data in unused sectors, such as S-tools, will take the file and spread the bits out throughout the free space on the floppy. Though undetectable in the normal Windows viewer, the file is still there.

Hidden Partitions

A hidden partition on a hard drive is another way of hiding information, sometimes large amounts, in plain sight. This method isn't terribly robust since a close examination of the hard drive, independent of the operating system, will usually reveal that the drive is bigger than the OS is letting on.

Slack Space

Slack space is a type of unused space in a disk. Even if the actual data being stored requires less storage than the cluster size, an entire cluster is reserved for the file. The unused space is called the *slack space*. If minimum space allocated is 32 kb and the file is 6 kb, then 26 kb is left unused and considered unavailable by the operating system. This unused space (slack space) could be used to hide information without showing up in any directory or file system.

Hiding in Network Packets

A covert channel is described as "any communication channel that can be exploited by a process to transfer information in a manner that violates the systems security policy." Essentially, it is a method of communication, referred to as Internet steganography, that is not part of an actual computer system design, but can still be used to transfer information to users or system processes that normally would not have access to the information. Hiding information within network packets can be accomplished a couple of ways. One is Packet Header Manipulation, which hides data in empty fields of the packet header, and Packet "sorting," which resorts packets in such a way that their sequence carries with it a meaning of either 1 or 0, which will ultimately yield a hidden message.

Issues in Information Hiding

Most things in life require trade-offs of some sort. Something for nothing isn't the way the universe works, and information hiding is no exception. The following is a list of limitations associated with trying to hide information steganographically.

Levels of Visibility

When looking at the first issue in information hiding, determining levels of visibility can be done by asking one question: Does the embedding process distort the cover to the point

where it is visually noticeable? If the image is unacceptably distorted, the carrier isn't sufficient for the payload; if it isn't distorted, the carrier is adequate.

Robustness vs. Payload

To have a robust method of embedding the message, you must have redundancy to resist changes made to the cover. This redundancy subsequently lowers the payload.

The exact opposite is also true. With little or no redundancy (robustness), the payload of the secret message can be larger.

more robust = lower payload

less robust = higher payload

File Format Dependence

Some image and sound files are either lossy or lossless. The conversion of lossless information to the compressed "lossy" information can destroy the hidden information in the cover. A good example is to think of the conversion of an uncompressed bitmap to a compressed, estimated jpeg. The compression and estimation change the bits to include the bits that might be the embedded message.

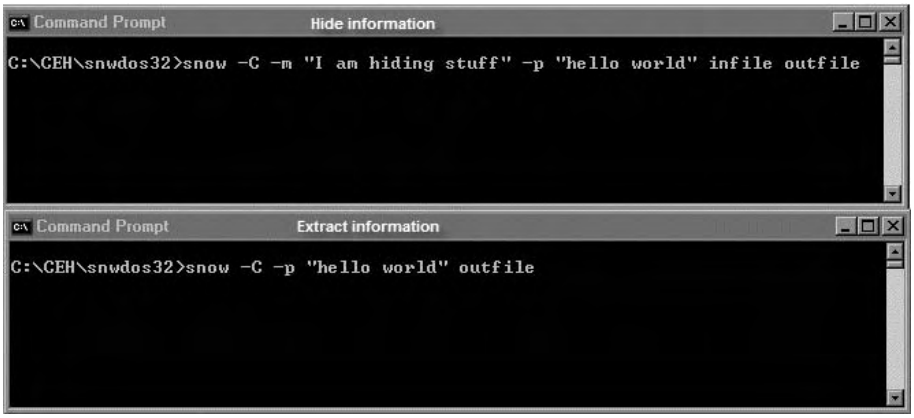
Steg Tools

There are a significant amount of tools available for creating steganographic content. These tools come in both open source and commercial versions. The majority of the tools can create steganographic content in a couple of file formats at best; this is because the tool creator wants to focus on a specific set of techniques and algorithms. You are best served to try the tools out, and see for yourself the capabilities of the tools. For this section we will discuss some of the popular tools that are used.

Snow

The tool Snow works by using whitespace to hide information. It is used to conceal text information by appending whitespace to the end of lines. Figure 18.1 shows the usage of Snow to hide a message and recover a message.

Figure 18.1 Hiding and Recovering a Message with Snow


Command Prompt
Hide information
C:\CEH\snwdos32>snow -C -m "I am hiding stuff" -p "hello world" infile outfile
Command Prompt
Extract information
C:\CEH\snwdos32>snow -C -p "hello world" outfile

Steganos

This tool has the capability to create steganographic content, and it can encrypt the information to provide another layer of security. It can hide files inside BMP, VOC, WAV, and text files.

Gifshuffle

This program has the capability of storing messages inside GIF images; it accomplishes this by shuffling the colormap. It can work with all GIF images, even the ones that contain transparency and animation. Another feature of the tool is it provides compression and encryption. It has two modes: message concealment and message extraction. Figure 18.2 shows the usage of Gifshuffle.

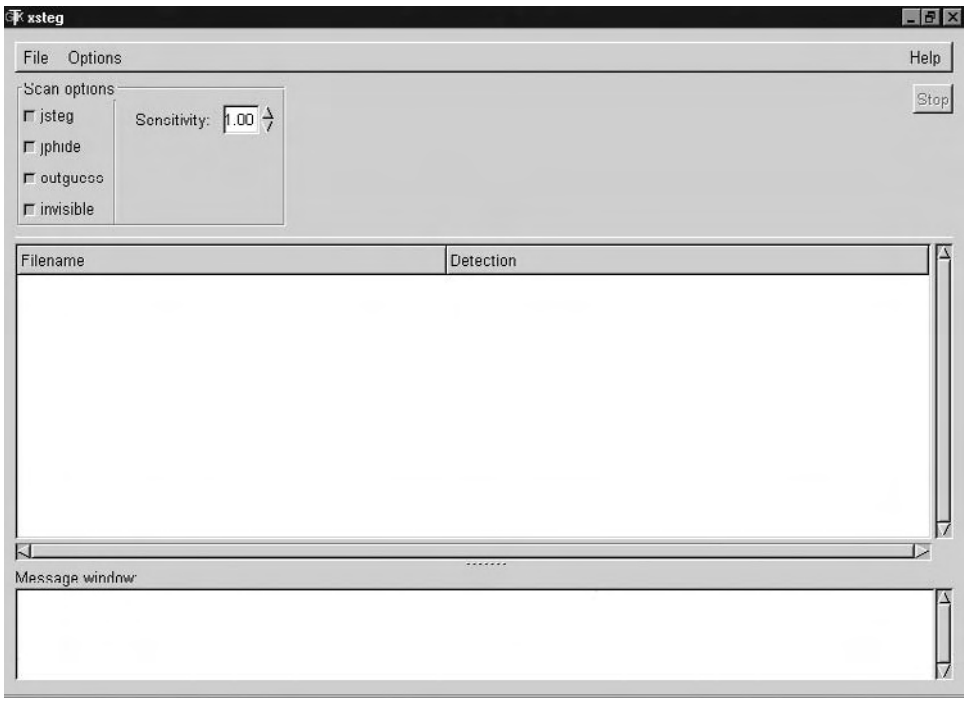
Figure 18.2 Using Gifshuffle


Command Prompt
Hide Information
C:\CEH\gswin>gifshuffle -C -m "meet me at 10" -p "hello world" in.gif out.gif
Command Prompt
Extract Information
C:\CEH\gswin>gifshuffle -C -p "hello world" out.gif

Outguess

Outguess is a powerful tool that allows the insertion of hidden information into the redundant bits of information found in our files. The tools supports PNM and JPEG file formats. Figure 18.3 shows a screenshot of the tool.

Figure 18.3 The Outguess Tool



Stegomagic

This tool will hide any kind of file or message, inside of a text file, WAV, or BMP. The tool can use either 24-bit, or BMP 256-bit color files. The data to be hidden does have a size limitation in that it can be approximately only one-eighth the size of the host or carrier file.

Table 18.1 contains a brief list of steganography and steganalysis tools. Study this list to get a quick understanding of the capabilities of these tools.

Table 18.1 Steganography and Steganalysis Tools

Tool	Description
Blindside	Hides files within bitmap (BMP) images
Camera/Shy	Works with Windows and Internet Explorer
	Uses GIF images
	Provides encryption and password protection
Cloak	Compresses and encrypts documents within images
Data Stash	Hides files within other files
	Password protection using Blowfish
Fort Knox	Uses MD5 and Blowfish algorithms
	Password protection capability
	Hiding and securing files and folders
Gifshuffle	Used to conceal messages in GIF images by shuffling the color map
	Works with all GIF images, including those with transparency and animation
Hermetic Stego	Hides a message file in BMP images
Hydan	Exploits redundancy in the i386 chipset
	Encodes information in machine code
	File size remains unchanged
	Message is Blowfish encrypted
Image Hide	Hides text in images
	Simple encryption capabilities
	Does not increase the file size
	Can evade mail sniffers
Invisible Secrets 4	Encrypts information for secure transfer across the network
	Can hide information in JPEG, PNG, BMP, HTML and WAV
JPHIDE and JPSEEK	Allows the hiding of a file in JPEG images
Masker	Provides encryption up to 448 bits
	Hides information in image, video, and sound files
Mp3Stego	Hides information in MP3 files during compressions
	Data is first compressed and encrypted and then hidden in the bit stream

Table 18.1 Continued

Tool	Description
Outguess	Inserts information into the redundant bits of data sources
	Supports PNM and JPEG formats
Pretty Good Envelope	Hides a binary message
	Appends a 4-byte pointer to the location of the start of the message
Snow	A whitespace program used to conceal messages in ASCII text by appending whitespace to the end of files
	Has built-in encryption capabilities
S-Tools	Hides multiple applications in a single object
Steganos	Combines two technologies: encryption and steganography
	Works with BMP, VOC, WAV, and ASCII files
StegaNote	Hides information in images or compressed files
Steghide	Allows compression of the hidden data and encryption of the embedded information
	Checksum integrity checking
	Support for JPEG, BMP, WAV, and AU files
Stegbreak	Steganalysis tool used to launch dictionary attacks against JPHIDE and Outguess
Stegdetect	Steganalysis tool that detects hidden information in JPEG files
Stego Suite	Contains several steganalysis tools
	Steg Watch can extract information using a dictionary attack.
	Stego Break was designed to crack a passphrase used in JPHIDE and JSteg.
StegSpy	A steganalysis tool that identifies the location of hidden content
	Identifies Hiderman, JPHIDE, Masker,
	JPegX, and Invisible Secrets
wbstego	Hides files in BMP images, HTML, PDF and text files

Steganography vs. Watermarking

It is common to think of *watermarking* when we refer to steganography; there are similarities, but there are also many differences. With steganography, the intent is to hide information; the information you decide to hide is hidden in its entirety. With the concept of watermarking, the intent is to protect the material from being used in violation of the creators intellectual property rights, and in some cases the watermark is clearly visible, and in some cases it is not. Whether or not the information is visible, the main objective remains the same; that is, protect the information from being copied and used without authorization. This usually involves some form of agreement with the creator, and most likely you will pay for the usage of the material.

Damage & Defense…

Digital Watermarking

Digimarc's Digital Watermarking tool allows a user to embed information into files. The information that is embedded is not visible during normal use, but it can be read by a computer or software. The information is imperceptible because the tool makes only subtle changes to the original content.

The definition of *watermarking* is "a form of marking that embeds copyright information about the artist or owner." In the traditional manufacture of paper, wet fiber is subjected to high pressure to expel the moisture. If the press's mold has a slight pattern, this pattern leaves an imprint in the paper—a watermark—which is best viewed under transmitted light. Today, the word "watermark" has been borrowed by high technology. Digital watermarks are imperceptible (or barely perceptible) transformations of digital data. Often, the digital data set is a digital multimedia object. Although digital images are most often mentioned in the same breath as digital watermarking, we note that watermarks can be applied to other forms of digital data, such as pictures, videos, and music. Now let's take a more in-depth look at the definition of, and methods used in, the technique of watermarking.

Fragile

The fragile watermark is one of the watermarking methods for authentication that has a low robustness toward modifications, where even a small change of the content will destroy embedded information, showing an attack attempt was made. In short, fragile watermarks are

supposed to break. Think of it as a digital version of the foil put under the cap of an aspirin bottle. It's pretty hard to get by that without it looking like it's been tampered with.

Robust

A robust watermark is almost exactly the opposite of a fragile watermark. A robust watermark can be either visible or invisible, depending on its purpose. Robust watermarks are very difficult to remove or damage. The easiest example of a robust watermark can be seen any time you turn on the television. The majority if not all channels have a watermark in the lower right-hand corner, not only to let you know what station you're on, but also to let you know that what you're watching is their content. Removing a visible watermark, especially in the television example, would be very difficult to do without someone noticing after the fact.

Attacking Watermarking

There are several known attacks that try to remove the watermark in a medium.

Mosaic Attack

The mosaic attack works by splitting an image into multiple pieces, and taking the pieces and putting them back together using javascript code. The web browser displays the information by putting the pieces back together and displaying them. The attack is successful, because in small pieces the watermark is unreadable.

2Mosaic

2Mosaic is a watermark breaking tool. The tool is an attack against a digital watermarking system. The attack is carried out by chopping an image into smaller pieces and creating an image piece. The perpetrator then embeds the images into a web page, the web browser renders the small image pieces that have been placed together, and it appears to be the real and original image without the watermark.

Detecting and Attacking Steganography

Steganography detection is usually a very difficult task, especially in instances of low payloads, since fewer bits are replaced, altered, or inserted. However, there are some ways to tell. The following is a list of methods, and a partial list of tools, used for detecting steganography. One final word for you on detection of steganography: there are a significant amount of false positives when trying to search through a possible steganography file; therefore, use caution when experimenting with the detection, and validate and verify all results you find.

Detection

When it comes to detection of steganography, there are several techniques that need to be discussed. As mentioned previously, when you are trying to detect steganography there are significant factors that have to be considered. We have several types of detection that we want to discuss in the following sections. These techniques are statistical tests, stegdetect, and stegbreak.

Statistical Tests

Statistical tests can reveal that an image has been modified by determining that its statistical properties deviate from the norm. Some tests are independent of the data format and just measure the entropy of the redundant data. Expect images with hidden data to have a higher entropy than those without.

Stegdetect

Stegdetect is an automated tool for detecting steganographic content in images. It is capable of detecting several different steganographic methods to embed hidden information in JPEG images.

Stegbreak

Stegbreak is a program that uses dictionary guessing to break the encoding password. It is used to launch dictionary attacks against JSteg-Shell, JPHide, and OutGuess.

Visible Noise

Attacks and analysis on hidden information may take several forms: detecting, extracting, and disabling or destroying hidden information. Images with too large a payload may display distortions from hidden information. Selecting the proper combination of steganography tools and carriers is important to successful information hiding.

Appended Spaces and "Invisible" Characters

Appended spaces and "invisible" characters refer back to the technique of hiding data in spaces within text.

This form of text semagram uses the white space in a document to denote binary values. The white space can be between individual words, sentences, or even between paragraphs. Almost any combination is possible, but to a point, if the text appears to have too much white space, it can be subject to scrutiny. Although this form of steganography can work effectively, it has a few big drawbacks. First, if the document is digital, any modern word processor would be able to show the spacing irregularities or, worse, reformat the document

and destroy the hidden information. The other drawback is that this method doesn't transmit a large amount of information easily, which can limit its practicality.

Color Palettes

Some tools have characteristics that are unique among stego tools. In some steganography programs the color palettes have unique characteristics that don't appear anywhere else. For example, the program Hide and Seek creates color palette entries that are divisible by 4 for all bit values. The palette modification creates a detectable steganography signature.

Attacking Steganography

It is usually easier to destroy than to create, and in the case of attacking steganographic files, this remains true. Since steganography is for the most part relatively fragile, a number of methods exist for attacking files that have steganographic content in them, or are suspected files that contain steganographic content. A couple of these may actually yield the hidden message after considerable work, but most simply destroy the message, making its recovery by anyone impossible. Furthermore, if the information is encrypted, then the best result is to destroy the carrier or host, because you are not going to break strong encryption without significant time and resources.

Steg-Only Attack

In this attack only the object that contains the steganographic information is available to the attacker. This is the most common situation; you would not leave the original host or carrier file available for someone to do an integrity check on.

Known-Cover Attack

In this attack you have gotten lucky, because the suspected perpetrator has not done a good job of covering their tracks. In this attack, the object with the stegonagraphic information and the original medium are both available. This makes it a trivial process to compare the original carrier or host file to the one with the hidden information.

Known-Message Attack

In this attack you know both the hidden message, and the object that has the steganographic message, and the process to complete the attack is simple.

Known-Stego Attack

When you conduct this type of attack, you have knowledge of the algorithm that is used to create the steganographic content, and the original file, and the file that has the hidden

information are available to you. Here we are trying to find the method used for embedding the content.

Chosen-Stego Attack

This attack is used when we have both the stego-medium and the steganography algorithm.

Chosen-Message Attack

The methodology of this attack is that the steganalyst generates an object containing steganographic information using some tool or algorithm of a chosen message, and looks for patterns in the created object that might lead to determining the actual tool or algorithm used to embed the steganographic content.

Disabling or Active Attacks

In this type of attack, the attackers change the carrier or host during the communication process. There are multiple techniques for doing this, as we will discuss now.

Blur

Smoothes transitions and decreases contrast by averaging the pixels next to the hard edges of defined lines and areas where there are significant color transitions.

Noise

Random noise inserts random colored pixels to an image, whereas uniform noise inserts pixels and colors that closely resemble the original pixels.

Noise Reduction

Reduces noise in the image by adjusting the colors and averaging pixel values.

Sharpen

Sharpen is the opposite of blur. It increases contrast between adjacent pixels where there are significant color contrasts, usually at the edge of objects.

Rotate

Rotation moves an image around its center point in a given plane.

Resample

Resampling involves an interpolation process to minimize the "raggedness" normally associated with expanding an image.

Soften

Soften applies a uniform blur to an image to smooth edges and reduce contrasts. It causes less distortion than blurring.

Application Password Cracking

Since the beginning of computer systems, there has always been some form of a password required for authentication into a system or network. This simple process has evolved in many ways, but it is still what we refer to as "user" controlled security. That is, the strength of this authentication method is largely in control of the user, because they are the ones selecting the password. For example, if you use a Personal Identification Number (PIN) and you select a simple PIN of all digits, that makes the PIN authentication system weak, and susceptible to cracking. The concept of cracking passwords is taking a password and decrypting it, or disabling the password protection of a system and/or network. In this section we will take a look at the common methods of attacking passwords, and we will close with some advice on how you can make your password more secure.

Types of Password Cracking

Since the first passwords were used, there have been methods to try and crack the actual text based version of the password. The reason we can crack passwords is two-fold. The users can select a weak password if the administrator has not enforced a strict password policy, and the other way is because the vendor has done a poor job with the scrambling of the password. There are several methods of attacking passwords that we will discuss. The methods are, guessing, dictionary, brute force, syllable attack, rule-based, and hybrid.

Guessing

In the guessing attack, perpetrators are successful when they are able to guess a person's password. This can be the result of selection by the user of a blank password. It can also be a result of choosing a simple password such as "password." Some users think they are smart, and will try a word in reverse like "drowssap." Another problem is when users select a password based on their kids, spouse, relative, or other personal information that is easy to identify.

Dictionary

With this attack you load a file of dictionary words into the password cracking tool, and if the password is one of the words within the dictionary file it is cracked. It is important to note that there are dictionary files available for many languages; therefore, it is a simple

process of loading your dictionary for the country you are conducting the testing in. Consultants have successfully cracked many passwords of foreign languages using this technique—a dictionary even exists for the Klingon language.

Brute Force

In the brute force method of password attacking, the concept is to try every possible combination of characters until a password is found. It is the slowest method of attack, but given enough time and resources it will discover any password.

Syllable Attack

This attack is a combination of brute force attack and dictionary attack. The technique usually is used when the password is known to be a nonexistent word.

Rule-Based

This technique is used when the perpetrator is able to get some information about the password, usually following some form of enumeration that has identified the password policy in place for an organization. For example, if the policy indicates the length of the password is not less than eight characters, and must contain at least numbers and a special character, then the perpetrator will adjust and customize the cracking tool for this.

Hybrid

A hybrid attack is used to find passwords that are a dictionary word with combinations of characters prepended or postpended to it. This attack is surprisingly successful, because in most cases users will select a password that is a dictionary word surrounded by additional characters.

Rainbow

The rainbow attack technique works by calculating all the possible hashes for a character set, and storing them in a table. The password hash is presented to the tool that uses the rainbow algorithm, and a table search is made until the password is found. This is a much quicker method that the other types of attack; however, the limitation of the rainbow technique is the size requirements for a table, so you need to think in the terms of terabytes for complex passwords.

Damage & Defense…

When You Cannot Crack the Password!

There are cases when no matter no hard you try you cannot crack the password, as more organizations implement a strong password policy, and eliminate the storage of the weak password hashes. The cumulative result of all this is, there are cases when you are not going to get the password cracked. So, what do we do? Quit? Even though you might not be able to crack the operating system password there are alternatives.

You could plant a keystroke logger, and wait for the user to login and then capture the password. The challenge of this is that in some cases, a keystroke logger will get you into possible legal trouble; therefore, you should seek legal counsel before you consider anything like this.

Another less intrusive way is to use a bootable media device like a CD-ROM or DVD, and boot from the media. Once you do this you can either change or blank the Administrator password account. It is usually easiest, and works better to blank the password. Most tools have this capability; all you need is access to the machine, and a BIOS that allows the boot from your chosen media device. We commonly do this for our own machines when we forget the password. With the auto-login that XP and other operating systems are defaulting to, it is more common that you or someone you know will forget a password. When this happens, insert the CD in and reboot, and reset or blank the password and you're in!

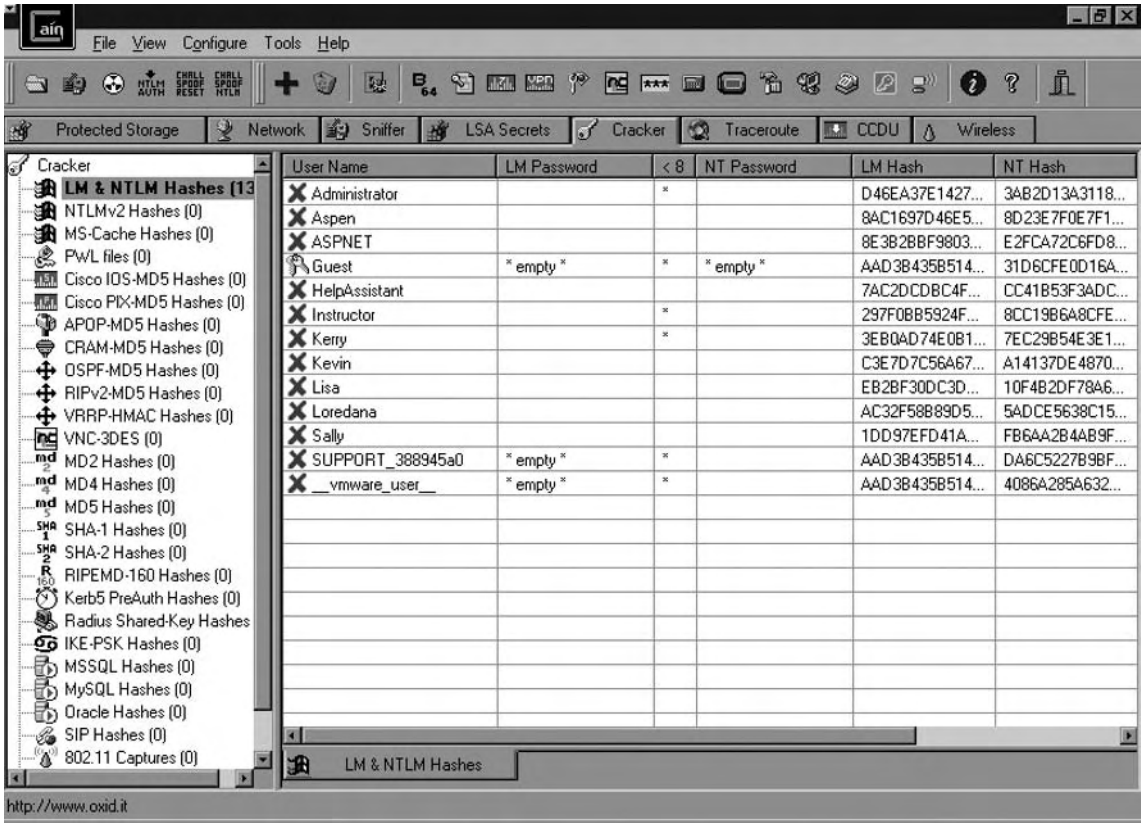
Password-Cracking Tools

Now that we have discussed the types of password attacks, it is time to discuss some of the tools. When it comes to cracking passwords, there are an extraordinary amount of tools that are available, so we have selected several of the most popular ones. The tools we have selected and will discuss next are cain and abel, lcp, ophcrack, john the ripper, brutus, and rock xp.

Cain and Abel

This is a Windows-based password recovery tool. It uses multiple methods to capture the password hashes. It can get the hash from the network, or dump it from the local machine. Cain and Abel uses dictionary attacks, brute force, and other cryptanalysis techniques to crack the password. This tool is much more than just a password cracking tool. The screen shot in Figure 18.4 shows the password cracking capabilities of this excellent tool.

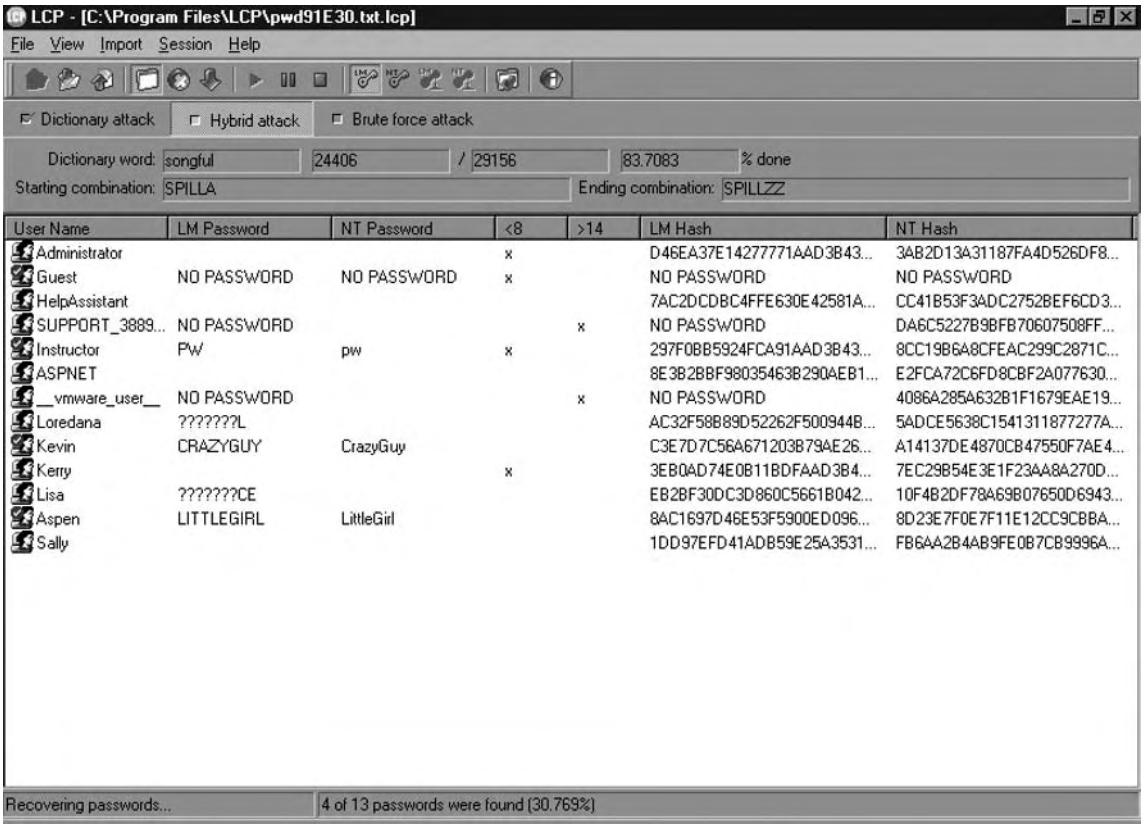
Figure 18.4 Password Cracking with Cain and Abel


File View Configure Tools Help
Protected Storage
Network
Sniffer
LSA Secrets
Cracker
Traceroute
CCDU
Wireless
Cracker
LM & NTLM Hashes (13
NTLMv2 Hashes (0)
MS-Cache Hashes (0)
PWL files (0)
Cisco IOS-MD5 Hashes (0)
Cisco PIX-MD5 Hashes (0)
APOP-MD5 Hashes (0)
CRAM-MD5 Hashes (0)
OSPF-MD5 Hashes (0)
RIPv2-MD5 Hashes (0)
VRRP-HMAC Hashes (0)
VNC-3DES (0)
MD2 Hashes (0)
MD4 Hashes (0)
MD5 Hashes (0)
SHA-1 Hashes (0)
SHA-2 Hashes (0)
RIPEMD-160 Hashes (0)
Kerb5 PreAuth Hashes (0)
Radius Shared-Key Hashes
IKE-PSK Hashes (0)
MSSQL Hashes (0)
MySQL Hashes (0)
Oracle Hashes (0)
SIP Hashes (0)
802.11 Captures (0)
User Name | LM Password | < 8 | NT Password | LM Hash | NT Hash
Administrator | | * | | D46EA37E1427... | 3AB2D13A3118...
Aspen | | | | 8AC1697D46E5... | 8D23E7F0E7F1...
ASPNET | | | | 8E3B2BBF9803... | E2FCA72C6FD8...
Guest | * empty * | * | * empty * | AAD3B435B514... | 31D6CFE0D16A...
HelpAssistant | | | | 7AC2DCDBC4F... | CC41B53F3ADC...
Instructor | | * | | 297F0BB5924F... | 8CC19B6A8CFE...
Kerry | | * | | 3EB0AD74E0B1... | 7EC29B54E3E1...
Kevin | | | | C3E7D7C56A67... | A14137DE4870...
Lisa | | | | EB2BF30DC3D... | 10F4B2DF78A6...
Loredana | | | | AC32F58B89D5... | 5ADCE5638C15...
Sally | | | | 1DD97EFD41A... | FB6AA2B4AB9F...
SUPPORT_388945a0 | * empty * | * | | AAD3B435B514... | DA6C5227B9BF...
__vmware_user__ | * empty * | * | | AAD3B435B514... | 4086A285A632...
LM & NTLM Hashes
http://www.oxid.it

LCP

The LCP tool was developed as a free alternative to the very popular L0phtcrack tool that was the pioneer in cracking passwords on a Windows platform. L0phtcrack is no longer offered, and LCP is an excellent way to get the features that used to be available with l0phtcrack. The tool offers the ability to import from a variety of formats, and uses dictionary, hybrid, and brute force attack methodologies to discover the password. Figure 18.5 shows the LCP tool with a dump of the SAM database.

Figure 18.5 LCP with a Dump of the SAM Database


LCP - [C:\Program Files\LCP\pwd91E30.txt.lcp]
File View Import Session Help
Dictionary attack
Hybrid attack
Brute force attack
Dictionary word: songful
24406
/ 29156
83.7083
% done
Starting combination: SPILLA
Ending combination: SPILLZZ
User Name | LM Password | NT Password | <8 | >14 | LM Hash | NT Hash
Administrator | | | x | | D46EA37E14277771AAD3B43... | 3AB2D13A31187FA4D526DF8...
Guest | NO PASSWORD | NO PASSWORD | x | | NO PASSWORD | NO PASSWORD
HelpAssistant | | | | | 7AC2DCDBC4FFE630E42581A... | CC41B53F3ADC2752BEF6CD3...
SUPPORT_3889... | NO PASSWORD | | | x | NO PASSWORD | DA6C5227B9BFB70607508FF...
Instructor | PW | pw | x | | 297F0BB5924FCA91AAD3B43... | 8CC19B6A8CFEAC299C2871C...
ASPNET | | | | | 8E3B2BBF98035463B290AEB1... | E2FCA72C6FD8CBF2A077630...
__vmware_user__ | NO PASSWORD | | | x | NO PASSWORD | 4086A285A632B1F1679EAE19...
Loredana | ???????L | | | | AC32F58B89D52262F500944B... | 5ADCE5638C1541311877277A...
Kevin | CRAZYGUY | CrazyGuy | | | C3E7D7C56A671203B79AE26... | A14137DE4870CB47550F7AE4...
Kerry | | | x | | 3EB0AD74E0B11BDFAAD3B4... | 7EC29B54E3E1F23AA8A270D...
Lisa | ????????CE | | | | EB2BF30DC3D860C5661B042... | 10F4B2DF78A69B07650D6943...
Aspen | LITTLEGIRL | LittleGirl | | | 8AC1697D46E53F5900ED096... | 8D23E7F0E7F11E12CC9CBBA...
Sally | | | | | 1DD97EFD41ADB59E25A3531... | FB6AA2B4AB9FE0B7CB9996A...
Recovering passwords...
4 of 13 passwords were found (30.769%)

Ophcrack

Ophcrack is a Windows-based password cracker that uses the concept of the rainbow cracking methodology by conducting the crack from existing rainbow tables. The algorithm deployed is based on the time-memory trade-off technique of precomputing all possible hashes and then applying the hash to the table.

John the Ripper

John the Ripper (JTR) is a fast password cracking tool that will not only crack Windows-based passwords, but also passwords on Unix and Linux systems. The tool runs both within a Unix and a Linux environment.

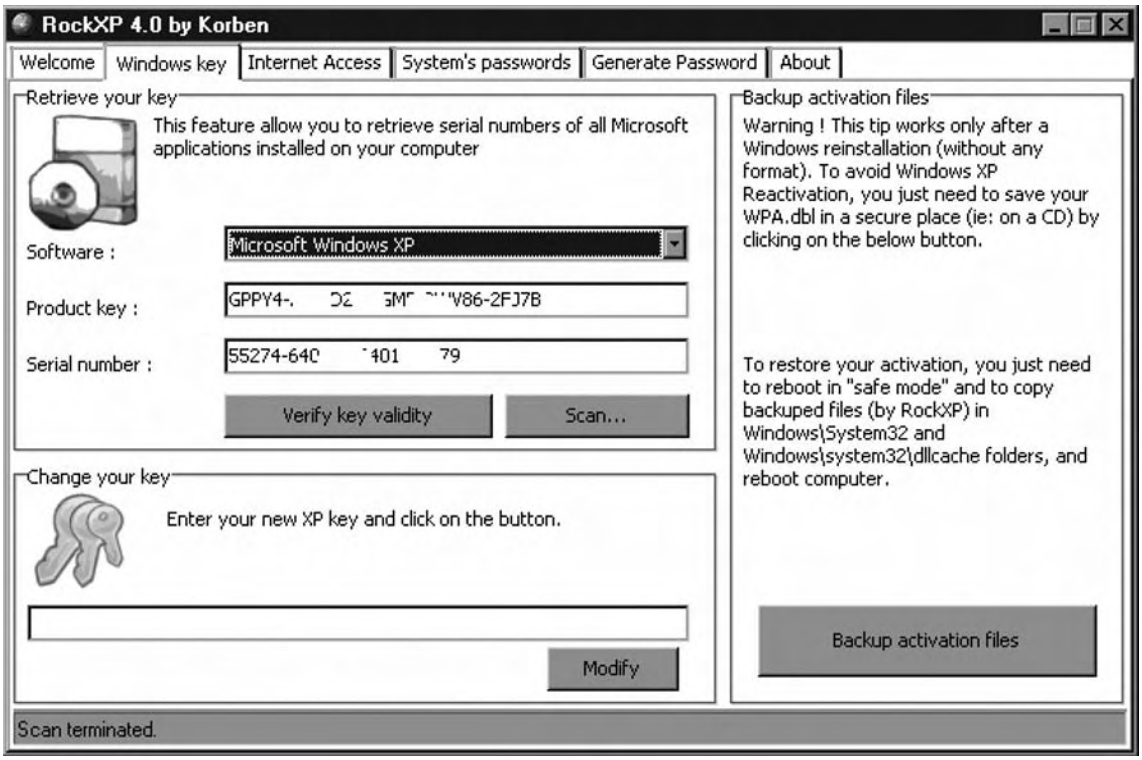
Brutus

Brutus is a very fast and flexible password cracking tool that can perform the cracks remotely. It commonly is used to crack Web site passwords. It is a Windows-based tool that can support up to 60 simultaneous target connections.

Rock XP

Have you ever tried to install a copy of Windows, or a Microsoft product, but could not find the product key to do the installation? Well you are in luck—this tool excels at recovering that lost information. The tool allows you to recover your XP key, save your XP activation, and even change your XP key! Figure 18.6 shows the recovery of an XP product key.

Figure 18.6 Recovering an XP Product Key with Rock XP


RockXP 4.0 by Korben
Welcome
Windows key
Internet Access
System's passwords
Generate Password
About
Retrieve your key
This feature allow you to retrieve serial numbers of all Microsoft applications installed on your computer
Software :
Microsoft Windows XP
Product key :
Serial number :
Verify key validity
Scan...
Change your key
Enter your new XP key and click on the button.
Modify
Backup activation files
Warning ! This tip works only after a Windows reinstallation (without any format). To avoid Windows XP Reactivation, you just need to save your WPA.dbl in a secure place (ie: on a CD) by clicking on the below button.
To restore your activation, you just need to reboot in "safe mode" and to copy backuped files (by RockXP) in Windows\System32 and Windows\system32\dllcache folders, and reboot computer.
Backup activation files
Scan terminated.

Common Recommendations for Improving Passwords

As we discussed earlier, the users can assist us when it comes to making the passwords more difficult to crack. Now it is time to discuss some methods of improving your password security. There are some common-sense things to remember. The first is you cannot have a password that is a dictionary word, or a derivative of a dictionary word. There are several more we will discuss below such as no personal data and multiple character sets.

No Dictionary Words

As we just identified, the dictionary attack succeeds because users choose words that are found in the dictionary. Furthermore, there are attacks that use the hybrid approach, and append characters to the beginning or the end of a dictionary word. The safest password is one that has no resemblance to a dictionary word.

No Personal Data

As was shown in the popular 1980s movie *War Games,* a password based on the name of a child or a spouse is very weak. Additionally, a password should never be used that is known personal data like any type of identity number.

Multiple Character Sets

A password that is eight characters long, and is only alphanumeric is much easier to crack than a password that is eight characters long, but contains at least two additional character sets.

Do Not Store Weak Hashes

The Lan Man (LM) hash used in Windows systems has been shown to be susceptible to rapid cracking, and should not be stored on the machines or networks of today. There are options you can set in the Group Policy Object that will prevent the storage of these weak hashes. On Windows Server 2003, the LM hash is not stored by default.

Standard Password Advice

There are some things that you can do, and should do on a regular basis, to improve your password security. The following section highlights the recommended standards to maintain when maintaining and administering your passwords. These include changing the password, no password reuse, and selecting the right size.

Change

The most important thing is to change your password periodically, because the longer a password is around the more susceptible it is to cracking.

Not in More Than One Place

A common mistake is to use the same password in multiple areas; this is a definite "no-no" as once a perpetrator obtains the password they will have multiple access to different places. Another thing to remember is when you use protocols that are in clear text, anyone can capture your password on the network, so protocols such as Post Office Protocol (POP) reveal the password to a network capture utility; therefore, it is imperative that these passwords be unique and used only in that one place.

Size

When it comes to passwords, the size is very important, because the longer the password, the more time it will take to crack it.

Creation

One technique for creating a password is to use a phrase or saying that you know well. A popular technique is the nursery rhyme, "Twinkle, Twinkle, Little Star, How I Wonder How You Are." In this technique we take the song, and the first letter of each word, and form our password. For example, Ttl*hIwhur!. Of note, when we use more than 14 characters the LM hash is not stored.

Summary

The chapter started with a look at the concept of steganography, and how it was defined as hiding information in plain sight. We continued the discussion with a recap of the long history of the usage of steganography, dating all the way back to the Greeks in ancient times.

Once we had covered the history of steganography, we discussed the categories of steganography and how it relates to forensics. We followed this discussion of steganography usage with the concept of detecting steganography. Once steganography has been used to hide information it can be difficult to detect. The detection is made much easier if the original file that was used to hide the information is available; however, as discussed this is rarely the case unless we are dealing with a novice hacker.

We also discussed the difference between encryption and steganography. With encryption we can see the encrypted information or data, but in steganography we are actually attempting to hide all information from view. It is important to note that some tools will combine encryption with steganography and the information or data that is hidden will still require some form of a key to allow the user to view it.

Another comparison we made was the comparison of a watermark with steganography. With a watermark we are applying a mark or code to our material to protect if from being copied illegally and in violation of intellectual property rights.

We continued in the chapter talking about passwords, and how they are attacked. We discussed the main types of password attacks that are commonly used against us. We looked at the techniques that are used within tools, and also the techniques that were of a nontechnical nature.

Our ability to crack the password relies on several factors, and we discussed how the user can select an easy-to-guess password, which would make it much easier for us to crack the password. We also discussed how if the administrator implemented strong password policies it makes cracking the password more difficult.

Within the password section we discussed that by default some operating systems store a weak version of the hash, or scrambling of the password, and when this is present it makes it much easier to crack the password.

We concluded this chapter with a discussion of proven strong password guidance, and additional guidelines for achieving a strong password policy, standard and baseline.

Chapter 19

PDA and Blackberry

Solutions in this chapter:

- **PDA Forensics**
- **Investigative Methods of PDA Forensics**
- **PDA Investigative Tips**
- **Deploying PDA Forensic Tools**
- **Introduction to the Blackberry**
- **Operating System of the Blackberry**
- **Blackberry Operation and Security Capabilities**
- **Forensic Examination of a Blackberry**
- **Attacking the Blackberry**
- **Securing the Blackberry**

☑ **Summary**

☑ **Frequently Asked Questions**

Introduction

In this chapter we will discuss the concept of conducting a forensic investigation on data that has been read, stored, or manipulated on some type of mobile device. The techniques for investigating a mobile device are similar to that of the more traditional storage devices; however, there are some notable differences that we need to be aware of while collecting potential evidence. (See Chapter 20 for a detailed discussion of the iPod.)

PDA Background Information

A Personal Digital Assistant (PDA) is a handheld computing device that combines a multitude of functions and features including computing, telephone, fax, and Internet. Additionally, the PDA can and most often contains some form of networking or other form of connectivity capabilities. Today, a PDA is a powerful device that can function as a cellular phone, a fax sender, a Web browser, and a personal organizer. These devices have reached such a level of power and functionality that they are in essence a mini-computer.

Components of a PDA

The PDA device has several components. Our intent here is to discuss some of the more common ones. The first component of the PDA is the microprocessor, which is similar to any other microprocessor except that there is a restriction on its size. Another component of the PDA is some form of input device, such as a touch screen. In addition to these components, an essential component is the operating system that is running the software for the PDA device.

PDA Forensics

As discussed previously, the concept of PDA forensics is very similar to the procedures and methodologies that are used with any form of forensics. When we discuss PDA forensics, there are investigative methods that you should use when performing a forensic investigation of a PDA.

Investigative Methods

There are four main steps when it comes to performing a forensic investigation of a PDA. These four steps are:

1. Examination
2. Identification
3. Collection
4. Documentation

We start off by securing the evidence. It is essential that you follow a process that has been approved by legal counsel to secure the PDA. When you seize the PDA you have to ensure that you take the PDA, the docking cradle, and the external memory cards. This is probably one of the most difficult things to control and requires that you conduct a thorough search for any and all memory cards. With the size of memory cards today, there is an extensive amount of evidence that you would be missing if you missed just one memory card. Once you secure the evidence, the next step is to create an exact image to preserve the crime scene. Once you have acquired the image it is time to examine the evidence. Once you have examined the evidence, you have to present it, which is usually done by compiling an extensive report based on the investigation thus far. It is also your responsibility as the examiner to maintain the evidence, which consists of keeping it in a secure location. You also have to ensure that the PDA remains charged so that the data and information is maintained in a constant state.

Step 1: Examination

In the examination step of PDA forensics, you first need to understand the potential sources of the evidence, which can be the device, the device cradle, the power supply, and any other peripherals or media that the device being examined has came into contact with. In addition to these sources, you should also investigate any device that has synchronized with the PDA you are examining.

Step 2: Identification

In the identification step of PDA forensics, you start the process by identifying the type of device you are investigating. Once you have identified the device, you then have to identify the operating system that the device is using. It is critical to the investigative process that you determine the operating system. Furthermore, once you have identified the operating system, it is important to note that it is possible that the device could be running two operating systems. During the identification process, there are several interfaces that can assist you including the cradle interface, the manufacturer serial number, the cradle type, and the power supply itself.

Step 3: Collection

During this part of the forensic investigation, it is imperative that you collect data and potential evidence from the memory devices that are part of or suspected to be part of the PDA you are investigating. There are a multitude of these types of devices, so we will limit our discussion to just a few, including the SD, the MMC semiconductor cards, the

micro-drives, and the universal serial bus (USB) tokens. The SD cards range in size from a few megabytes (MB) to several gigabytes (GB), and the USB tokens can range from a few MBs to multiple GBs. In addition to seizing and collecting the memory devices, you also have to collect the power leads, cables, and any cradles that exist for the PDA. Extending the investigation process further, it is imperative that you collect all of the types of information consisting of both volatile and dynamic information. Consequently, it is imperative that you give the volatile information priority while you collect evidence. The reason for giving this information priority is because anything that is classified as volatile information will not survive if the machine is powered off or reset. Once the information has been captured, it is imperative that the PDA be placed into an evidence bag and maintained at stable power support throughout.

Step 4: Documentation

As with any component in the forensic process, it is critical that you maintain your documentation and "chain of custody." As you collect information and potential evidence, you need to record all visible data. Your records must document the case number and the date and time it was collected. Additionally, the entire investigation area needs to be photographed, which includes any devices that can be connected to the PDA or currently are connected to the PDA. Another part of the documentation process is to generate a report that consists of the detailed information that describes the entire forensic process that you are performing. Within this report you need to annotate the state and status of the device in question during your collection process. The final step of the collection process consists of accumulating all of the information and storing it in a secure and safe location.

PDA Investigative Tips

When it comes to the PDA device, there are several things you need to consider while carrying out an investigation These devices can be managed and maintained at all times. Adding further complication is the fact that with PDA devices, a suspect can have immediate access 24 hours a day, 7 days a week. Another thing that makes your job as an investigator more challenging, is that PDAs are immediate boot cycle devices. Accordingly, it is important to remember these devices typically contain a plethora of information and are a vault of evidence for the forensic examiner.

Device Switched On

When you are beginning your investigation process, and discover that the PDA that you want to process for evidence is in the "on" mode, it is imperative that you act immediately

and get power to the PDA, so that it will not lose the volatile information that could quite possibly be essential to your evidence collection process.

Device Switched Off

If the device is in the "off" state, you should leave the device in this state and then switch the device on and take a picture of it. Additionally, you need to note and record the current battery charge.

Device in Its Cradle

Avoid any further communication activities with the device and remove any connections from the PC device. It is important to note that there is a possibility that a sophisticated suspect might have a "tripwire" device, and once you disconnect the PC this could activate the device, which in turn could run a script that might erase potential evidence. Despite this possibility, you have to disconnect the device to continue the investigation.

Device Not in Its Cradle

If the device is not in the cradle, your investigative requirements are made much simpler, because there is no danger of a "tripwire" being triggered. With the device being out of its cradle, you simply seize the cradle and any cords associated with it.

Wireless Connection

Avoid any further communication activities if possible. Eliminate any wireless activity by placing the device into an envelope that can isolate it. This envelope also needs to provide anti-static protection so that the device is not damaged.

Expansion Card in Slot

Do not initiate any contact that requires taking components off of the device, or that requires you to open the device in any way. This includes any and all peripheral devices and/or media types of cards.

Expansion Sleeve Removed

The first thing to accomplish is to seize the sleeve itself, and also seize any and all related peripherals and media cards.

Notes from the Underground…

Impact of Mishandling PDA Devices

While conducting an investigation of a potential crime scene, our team discovered a sticky note with a password written on it. The team member entered the password, but it did not work, so he continued to try to get access, but after the tenth attempt the Blackberry did a complete data wipe, and whatever information was on that device was lost. This is because there is software that will log the attempts at entry and will do a complete wipe after a certain amount of invalid login attempts.

Deploying PDA Forensic Tools

When you are conducting a forensic investigation, there is no shortage of tools available. However, when investigating handheld or PDA devices, there are not as many tools as a typical forensic investigator has.

PDA Secure

The first tool to discuss is the PDA Secure tool. This tool offers enhanced password protection, along with encryption, device locking, and data wiping. The PDA Secure tool allows administrators greater control over how handheld devices are used on networks. Additionally, it allows you to set a time and date range to monitor information such as network login traffic, infrared transmissions, and any applications being used.

PDA Seizure

PDA Seizure is a comprehensive tool that assists in seizing the PDA. It allows the data to be acquired, viewed, and reported on. PDA Seizure works within a Windows environment and can extract the random access memory (RAM) and read only memory (ROM). It has an easy-to-use graphical user interface (GUI), and includes the tools that are needed to investigate the files that are contained within the PDA.

PDA Seizure provides multi-platform support, where the forensic examiner can acquire and examine information on PDAs for both the Pocket PC and Palm operating system (OS) platforms. The PDA Seizure tool has a significant amount of features, including forensic imaging tools, searches on data within acquired files, hashing for integrity protection of acquired files, and book-marking capability to assist the examiner in the organization of information.

EnCase

EnCase is one of the most popular commercial forensic tools available, and can be used to acquire information and evidence from a PDA. The EnCase tool can acquire images, and also consists of tools that allow you to conduct complex investigations efficiently and accurately.

Introduction to the Blackberry

The Blackberry is also known as a RIM device and is equipped with the RIM software implementation of proprietary wireless-oriented protocols. Furthermore, the device is supported by the RIM Blackberry Message Center. The Blackberry shares similarities with the PDA devices discussed earlier; however, it is always on and participating in some form of wireless push technology. As a result of this, the Blackberry does not require some form of desktop synchronization like the PDA does. This unique component of the Blackberry adds a different dimension to the process of forensic examination, and in essence this portability can be the examiners greatest ally.

Operating System of the Blackberry

The current version of the Blackberry OS has numerous capabilities and features, including over-the-air activation, the ability to synchronize contacts and appointments with Microsoft Outlook, a password keeper program to store sensitive information, and the ability to customize your Blackberry display data.

Blackberry Operation and Security

The Blackberry device has an integrated wireless modem, which allows it to communicate over the BellSouth Intelligent Wireless Network. The Blackberry uses the Blackberry Serial Protocol, which is used to backup, restore, and synchronize the data that is communicated between the Blackberry handheld unit and the desktop software. This protocol comprises simple packets and single byte return codes. The device uses a strong encryption scheme that safeguards confidentiality and authenticity of data. It keeps data encrypted while in transit between the enterprise server and the device itself.

Wireless Security

The Blackberry has a couple of transport encryption options, which are the Triple Data Encryption Standard (DES) and the Advanced Encryption Standard (AES). Those who want to implement the most secure method will elect to encrypt with the AES algorithm. The Blackberry has another feature that is referred to as the Password Keeper, which

offers the capability of securely storing password entries on the devices, which could consist of banking passwords, PINs, and so on. This important information is protected by AES encryption.

Security for Stored Data

There are several capabilities available on the Blackberry when it comes to securing the data that is stored there. The first option we will discuss is the capability to make password authentication mandatory through the customizable Information Technology (IT) policies on the Blackberry Enterprise Server. An additional method of protection from unauthorized parties is the fact that there is no staging of data between the server and the Blackberry where data is decrypted.

Forensic Examination of a Blackberry

Since the Blackberry is an always-on, push messaging device, information can be pushed to it at any time. It is important to note that the information that is pushed has the potential of overwriting any data that was previously deleted. The problem is compounded by the fact that, without warning, there are a multitude of applications that may receive information and make the attempts by the forensic investigator to recover information and an unaltered file system much more difficult. The first step in preserving the information is to eliminate the ability of the device to receive this data push. If possible, you can turn the radio off, or a better solution is to take the device to an area where the signal cannot be received, such as putting the device inside of a filing cabinet drawer. One might think, "I'll just turn it off." This would be a serious mistake! The Blackberry is not really "off" unless power is removed for an extended period, or the unit is placed in storage mode. Furthermore, once the unit is powered back on, any items that were in the queue waiting to be pushed to the device could possibly be pushed before you could stop them. As mentioned previously, it is quite possible that a change to state such as a power-off of the Blackberry could result in a program being run on the unit that will allow the device to accept remote commands via e-mail.

Acquisition of Information Considerations

The considerations for the Blackberry are similar in some ways to the PDA devices, but there are some differences. Lets look at the considerations you have to make when acquiring evidence from the Blackberry.

Device is in the "Off" State

If the unit is off at the time of acquisition, the investigator needs to take it to a shielded location before attempting to switch it on. If a shielded location is not readily available, you might have success using a safe or other room that can block the signal well enough to

prevent the data push. One thing to consider is having a unit available that you can use to walk the network and look for weak coverage areas to use.

Device is in the "On" State

If the device you are examining is in the "on" state, then as outlined and detailed above, you need to take the device to a secure location and disable or turn off the radio before beginning the examination.

Password Protected

One thing to consider when it comes to password protection is the fact that the password itself is not stored on the device. The only thing that is stored on the device is a hashing of the plain text password. This storage is similar to the storage used by the majority of operating systems.

Evidence Collection

To collect evidence from the Blackberry you have to violate the traditional forensic methods by requiring the investigator to record logs kept on the unit that will be wiped after an image is taken. There are several different log files that you want to collect evidence from, including:

- **Radio Status** This log lets you enumerate the state of the device's radio functions.
- **Roam and Radio** This log has a buffer of up to 16 entries, records information concerning the tower, channel and so on, and will not survive a reset.
- **Transmit/Receive** This log records gateway information and the type and size of data transmitted.
- **Profile String** This log contains the negotiation with the last utilized radio tower.

Once the log information is extracted and enumerated, the image will be taken. If you do not require or need the log information, the image can be acquired immediately.

Unit Control Functions

The logs are reviewed by using the unit control functions. The first function is the Mobitex2 Radio Status, which provides information on the Radio Status, Roam and Radio Transmit, or Receive and Profile String. The second control function is the Device Status, which provides information on memory allocation, port status, file system allocation, and central processing unit (CPU) WatchPuppy. The third control function is the Battery Status, which

provides information on the battery type, load, status, and temperature. The last control function we will discuss is the Free Mem, which provides information on memory allocation, Common Port File System, WatchPuppy, OTA status, Halt, and Reset.

Imaging and Profiling

When you are conducting a forensic examination of a Blackberry, you need to conduct imaging and profiling. This is accomplished by extracting the logs from a developed image and acquiring an image of a bit-by-bit backup using the Blackberry Software Development Kit (SDK). The SDK is available from www.blackberry.com and is essential for the forensic examiner when investigating a Blackberry. The SDK utility dumps the contents of the Flash RAM into a file. Once the Flash RAM is dumped, it can be examined and reviewed using traditional methods with your favorite hex editor or other tool. In addition to reviewing the evidence with traditional methods, you can use the Simulator from the SDK to match the network and model of the investigated unit.

Attacking the Blackberry

There are several tools and methods available that allow you to attack a Blackberry, The first tool is the Blackberry Attack Toolkit, which along with the BBProxy software can be used to exploit Web site vulnerabilities. The second tool is the Attack Vector, which links and tricks users by downloading malicious software to the Blackberry. The last method we will discuss is the method of hijacks (or blackjacks). As the name implies, this allows someone to hijack a legal user's Blackberry and replace them on the network with potentially harmful devices.

Securing the Blackberry

There are several things you can do to secure the information on a Blackberry. The first thing you can do is clean the Blackberry memory, and protect stored messages on the messaging server. You can encrypt the application password as well as the storage of it on the Blackberry. Furthermore, you can protect storage of user data on a locked Blackberry by limiting the password authentication attempts. It is possible to set a maximum of 10 attempts to gain access to the device. Additionally, you can use AES technology to secure the storage of the password keeper and password entries on the Blackberry.

Information Hiding in a Blackberry

There are several places where you can hide information in a Blackberry. You can create hidden databases and hide information in partition gaps. Data can also be hidden in the gap between the OS/application and file partitions.

Blackberry Signing Authority Tool

The Signing Authority tool helps developers' protect data and intellectual property, and enables them to handle access to their sensitive Application Program Interfaces (APIs). The tool provides this protection by using public and private signature keys. It does this by using asymmetric cryptography to validate the authenticity of the request. Furthermore, the signing tool allows developers to exchange API information in a secure manner and environment.

Summary

The chapter started with an introduction to the PDA, and how the technology of today has pretty much provided us with a handheld computer. We continued the discussion with a look at the concept of PDA forensics and how many of the same things must be considered in forensics on normal systems. We also discussed some of the differences that must be considered when performing forensics on PDA devices.

We then discussed the methods of investigating a PDA. We talked about securing the evidence, and how the PDA, docking cradle, and any external memory cards should be seized. The next method we discussed was the acquiring of the evidence. We covered how you have to create an exact image of the evidence, and once the evidence is secured and acquired, the need to go on and examine the evidence that was acquired.

We then talked about the forensic examination considerations when confronted with a Blackberry (RIM) device. We concentrated on how the Blackberry has similarities to the PDA, but one way that they differ is that the Blackberry does not require synchronization to receive a significant amount of information. The Blackberry is always on, and to make the task a little more difficult, it is in a state where it is susceptible to receiving push technology updates at any time. Therefore, we discussed how it is imperative that we take this into account when preparing to examine the Blackberry. We also discussed the software that is available to assist us when we are examining the Blackberry. One excellent package of software is the SDK from Blackberry. We also discussed some of the tools available and the ways to attack the Blackberry, including the Blackberry Attack Toolkit, the Attack Vector, and the forms of hijacking (or blackjacking). Finally, we ended this chapter by discussing the methods of securing the Blackberry. We discussed the Blackberry Signing Authority Toolkit that provides tools to help developers protect their data and intellectual property, and uses asymmetric cryptography to authenticate information.

Frequently Asked Questions

Q: When conducting a forensic investigation of a PDA, what is the first step in the process?

A: As with any forensic examination, the first step is to have permission to seize the evidence that is required for your investigation.

Q: What sort of tools do I use to conduct a forensic examination of a PDA?

A: Most of the forensic tools that work with images will create an image of a PDA file system. The commercial software product EnCase has this capability, as do many others.

Q: If I am preparing to conduct an investigation of a PDA, why must I maintain the charge to the device?

A: Similar to a regular PC, the PDA device has both volatile and non-volatile information, and if the power is not maintained, there is a possibility you could lose information.

Q: Isn't a PDA and a Blackberry the same thing?

A: It is not uncommon to make this assumption, and there are similarities, but there are also many differences. The Blackberry is an always-on device that can be pushed information at any time, and unlike the PDA, the Blackberry does not require synchronization with a PC.

Q: How would I get access to log files on the Blackberry?

A: Some of the best tools for conducting an investigation of a Blackberry come from Blackberry themselves. There is a Software Development Toolkit (SDK) that can access and collect log files and other information.

Chapter 20

MP3 Forensics

Solutions in this chapter:

- **History**
- **Why Is an iPod Considered Alternative Media?**
- **Imaging and Hashing**
- **Hardware vs. Nonhardware Imaging**
- **Types of iPods**
- **File Systems**
- **"Hacking Tools" and Encrypted Home Directories**
- **Evidence: Normal vs. Not Normal**
- **Analysis Tools**

☑ **Summary**

Introduction

I remember when I got my first boombox, circa 1983. It was sea-foam green and had two tape decks and a radio. It came with batteries and was so small and portable that I could take it with me everywhere I went. Then there was my first Walkman, the size of a brick, with giant, spring-loaded headphones. And after that, my first portable CD player, which cost almost a month's wages back in 1990. Today, of course, portable music players are commonplace and mundane. Although five years ago having white iPod headphones on meant you were hip and stylish, now it just means you like to listen to tunes. But more and more, the iPod is becoming a medium to store not just music, as we will see.

At a conference, I had a conversation with a federal agent about a child pornography case that he had worked on. He and his colleagues had raided the suspect's home as usual, but noticed that the suspect was strangely unalarmed and even smirking as the agents copied all the data from his computers. He continuously proclaimed his innocence. The agent remembered a presentation I had given about iPods being used to store data and noticed that the suspect had an iPod on his desk. According to the warrant served on the suspect, the agents were permitted to seize all electronic equipment capable of storing data. The agent picked up the iPod, and suddenly the suspect's demeanor and attitude changed—he turned pale and became agitated. Before the raid was finished, the suspect had confessed that he routinely erased his computers' hard drives after transferring all of his child pornography photographs to his iPod. The evidence resulted in a conviction.

This chapter of the book will demonstrate how iPods can be used to store any type of data or information that can be stored on a regular personal computer. The data can be encrypted, hidden, and easily manipulated by the user. In this chapter, we will explore the ways to store, access, and find data stored on iPods, and the tricks people use to hide malicious data.

History

In the late 1990s, digital music began to gain in popularity. The MP3 music format was portable and the sound quality was closer to that of compact discs than the analog tapes that had been used for decades prior could ever be. The MP3 format used compression so that the files were small in size and could easily be stored and listened to on a personal computer.

Before the paint had even dried on MP3 technology people had figured out ways to cheat the system. The new MP3 technology led to two of the first file-transferring networks: Napster and Gnutella, which allowed users to share and download music without purchasing it, leading to government and retail industry outrage and uproar. This was a gray area at first, since there were no laws on the books about file transferring or sharing from peer-to-peer networks and individual users. Eventually Napster, Gnutella, and most other similar networks were shut down by government authority.

Soon after MP3 technology caught on, the first digital media players became available, but there was still not a commercially viable way to legally purchase and download MP3s until 2000–2001, when Apple released the iPod and its iTunes online retail download service. Suddenly consumers had a very good way to legally purchase music and transfer it to their iPods or other digital media players. The digital music industry boomed, and since then it has become a powerful segment of the music business as a whole.

Although there are many different brands and types of digital media devices, Apple's iPod accounts for more than 80 percent of the digital media market, so this chapter of the book will focus exclusively on the iPod.

Why Is an iPod Considered Alternative Media?

iPods have standard file systems of either Apple's HFS+ or Microsoft's FAT32, which we will explore in greater detail later. These file systems are static because they are not continually transferring data like other types of file systems, such as cell phones, for instance. Because of their static nature, performing forensics on iPods is not substantially different from performing forensics on a regular computer hard drive. The difference between an iPod and a regular computer that makes an iPod an alternative media device is that the primary function of an iPod is as a music player. Only recently have iPods evolved into photo storage and video player devices. Because they are used for entertainment purposes, iPods might not be thought of as data repositories containing evidence. What follows is a detailed description of iPod forensics and the process that my colleagues and I perform on iPods to extract and analyze data contained on them.

Imaging and Hashing

The first step in iPod forensics is to create an image of a device and hash it to ensure integrity. With digital evidence, we do not work on original evidence. Instead, we attempt to create a duplicate of the evidence. This duplicate can be an exact replica of all data contained on the device. There are two types of images: a physical image and a logical image. A *physical image* is a bit-for-bit copy of all data contained on a device, and a *logical image* is an image of the file system exactly as it appears on a device. Sometimes it may not be feasible to collect a physical image, and therefore a logical image is your only option. For example, you may have a warrant that will only allow you to copy a user's home directory. In that case, you would not be able to collect any data that was outside of the user's home directory.

For forensic purposes, a physical image is always the preferred type.

To preserve the integrity of the data, forensic examiners perform what is called a "hash" at every step of the way. A *hash* is a one-way mathematical algorithm that acts as a "fingerprint"

of all data contained on a device. This ensures that the data has not been altered from its original state at any point during the imaging process. Hashes can be performed by using tools such as md5sum. The tool is applied to a file and returns a number that corresponds to a particular algorithm. Then the imaging is performed and the md5sum tool is rerun. If any part of that file is altered after the md5sum tool is rerun, the number will change, signifying a potential loss of evidence integrity. This will ensure that the data you are working on has not been altered.

Another way to preserve the data is to use a write blocker. Write blocking a device will protect the device from any manipulation. It will essentially guard your evidence from being written to during imaging. This way, if you happen to make an error, your evidence will be protected. A write blocker typically comprises visible external hardware, such as Logibube's Forensic Talon or Intelligent Computing Solutions (ICS)'s Solo III. A hard drive is physically attached to one of these devices, which will ensure that no writes can be made to the evidence contained on that drive.

Since iPods do not have IDE interfaces and use FireWire or USB, they cannot be synced to standard IDE imaging devices such as the Solo III without a USB adapter. Similarly, an iPod should not be synced into a forensic tower and imaged using a Windows-based tool. When you plug any USB or FireWire device into a Windows machine, Windows will "touch" the device and change the files contained on the device. One of the most important rules that forensic examiners must follow is not to alter evidence, including date and time stamps on evidence.

Hardware vs. Nonhardware Imaging

You can image data using hardware and software operating systems. It is highly recommended that if you have a hardware device that has a USB and/or FireWire interface, you should image the data using the hardware device. Hardware devices write-block very well and they leave little room for error. An excellent hardware device is the Tableu Forensic USB Bridge, but others like it are available on the market. They are relatively inexpensive and user-friendly. A nonhardware solution, such as an operating system like Linux or DOS, can be configured to not automatically mount a device when imaging. Linux is not a substitution for a write blocker, and is susceptible to human error.

Removing the Hard Drive

It is possible to obtain an image by removing the hard drive from an iPod if the iPod does not use flash memory. Please see the section "Types of iPods," later in this chapter, to see which ones contain hard drives.

You can remove a hard drive from an iPod, but this is prohibitive because removing a hard drive could break the device.

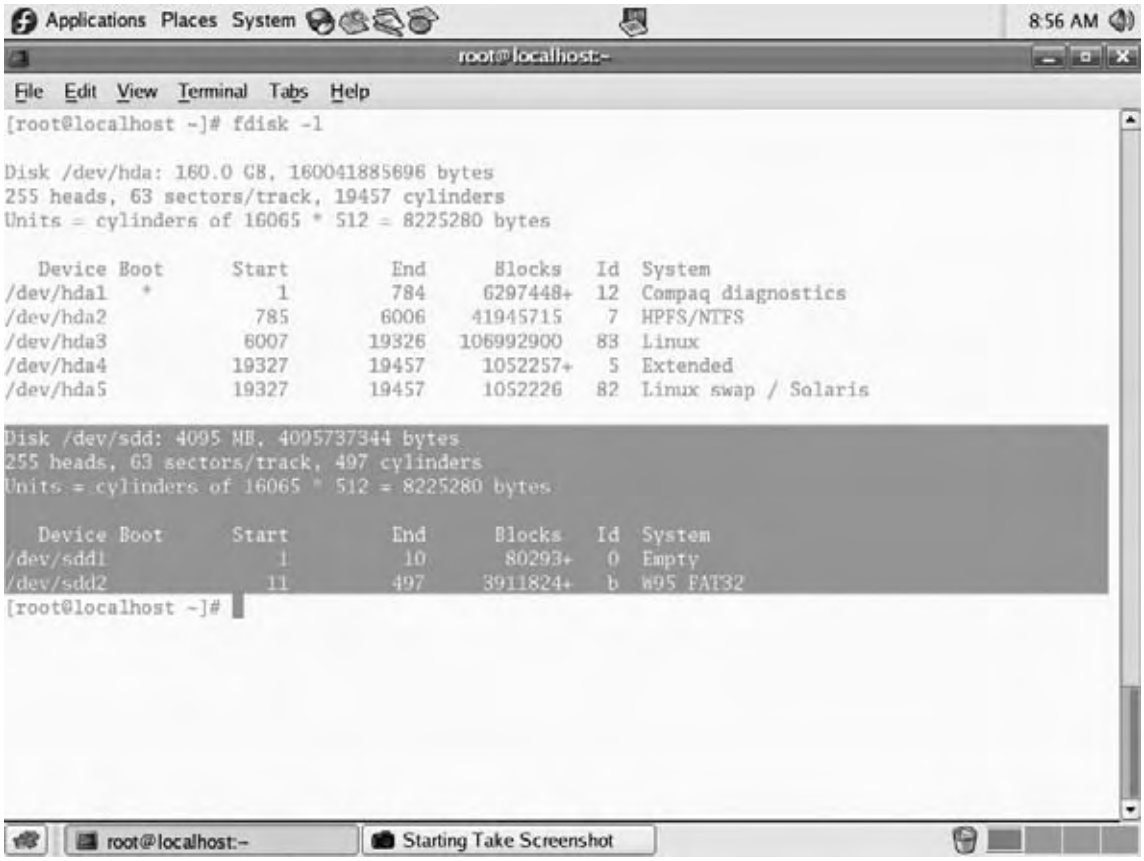
Linux

Another method you can use, but only if absolutely necessary, is to employ an operating system such as Linux for imaging. You can configure Linux to not automatically mount a USB device when the device is plugged into your forensic tower. This means that in theory, an iPod would remain untouched with files being unaffected when plugged in. This method allows for no write protection, however. If you make a mistake, you could destroy your evidence.

I suggest if you use this method that you employ the Linux command *dd* or the DCFL lab version, called *DCFLDD*, to image the device. The steps to perform this method follow.

The first thing that you see is the fdisk output of the device, with two partitions. In this case, the device /dev/sdd corresponds to the iPod device which is the target of the imaging process. The first entry in fdisk's output for /dev/had corresponds to the hard drive of the host computer used in the imaging operation and can be safely ignored (see Figure 20.1).

Figure 20.1 Fdisk Output


Applications Places System
8:56 AM
root@localhost:~
File Edit View Terminal Tabs Help
[root@localhost ~]# fdisk -l
Disk /dev/hda: 160.0 GB, 160041885696 bytes
255 heads, 63 sectors/track, 19457 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/hda1 * 1 784 6297448+ 12 Compaq diagnostics
/dev/hda2 785 6006 41945715 7 HPFS/NTFS
/dev/hda3 6007 19326 106992900 83 Linux
/dev/hda4 19327 19457 1052257+ 5 Extended
/dev/hda5 19327 19457 1052226 82 Linux swap / Solaris
Disk /dev/sdd: 4095 MB, 4095737344 bytes
255 heads, 63 sectors/track, 497 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Device Boot Start End Blocks Id System
/dev/sdd1 1 10 80293+ 0 Empty
/dev/sdd2 11 497 3911824+ b W95 FAT32
[root@localhost ~]#
root@localhost:~
Starting Take Screenshot

It is important to remember that the whole point of imaging in this way is to not mount the device. You can do everything you need to image the device without mounting.

The next step is to collect the MD5 hash of the device. You can perform this step in multiple ways, such as using another hashing tool or outputting an MD5 file to another directory. The following shows the command syntax for running the md5 checksumming utility "md5sum" on the target device "/dev/sdd" and storing the result in the file "/root/ipod.before.md5" (see Figure 20.2).

Figure 20.2 Collecting an MD5 Hash



Next, you view the /root/ipod.md5 file to make sure the hash is valid (see Figure 20.3). In forensics, it is good to double-check your work at every point, especially when there is no hardware write protection.

Figure 20.3 The "More" Command Displays the Contents of the File to the Screen



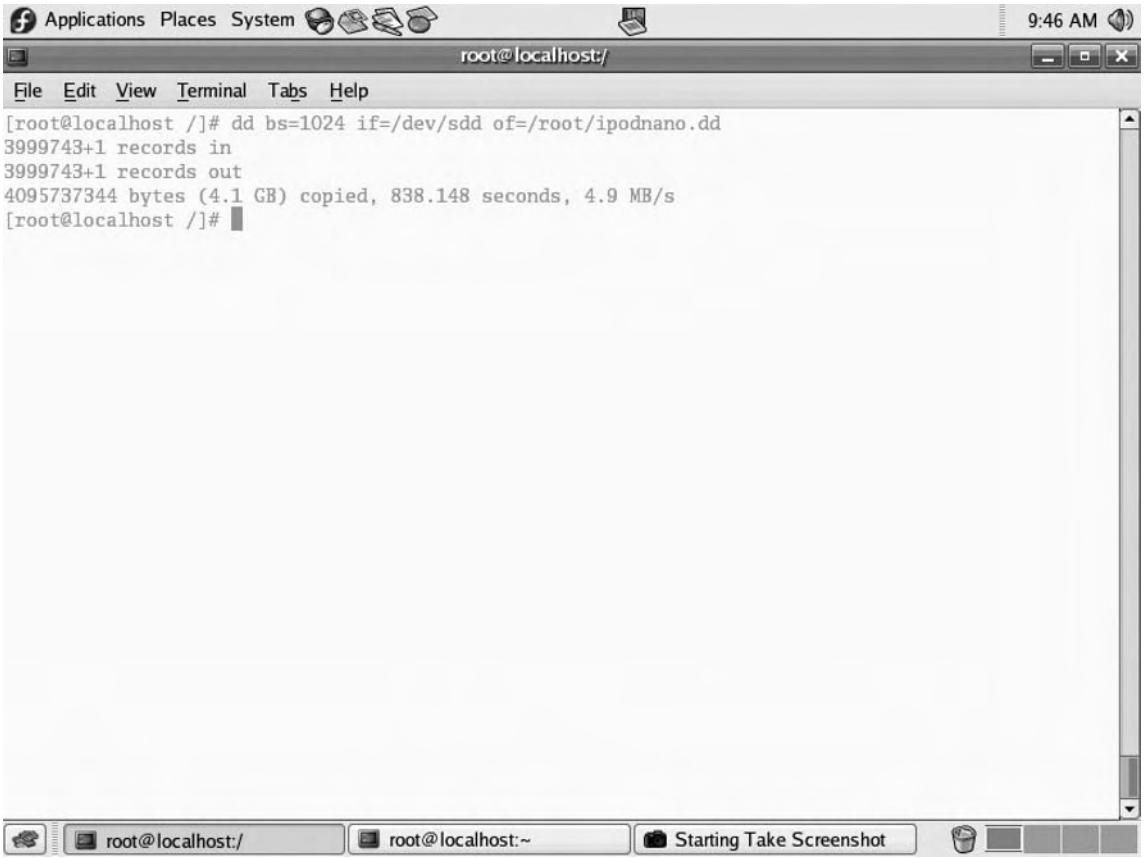
The next step is to create an image file from the device. This example uses the Linux "dd" command to image the data (see Figure 20.4). The BS option stands for "block size". Block size can change as desired, and has no impact on the data copied, except to optimize the throughput rate of the copy by copying that many bytes on each copy operation. The next two commands are the input file and the output file. It is important to double check that an iPod device is the input file and not the output file. Putting the iPod device as the "of" parameter could alter the contents of the evidence drive!

Figure 20.4 Imaging a Device



Figure 20.5 is an example of a completed *dd* function.

Figure 20.5 A Completed dd Function



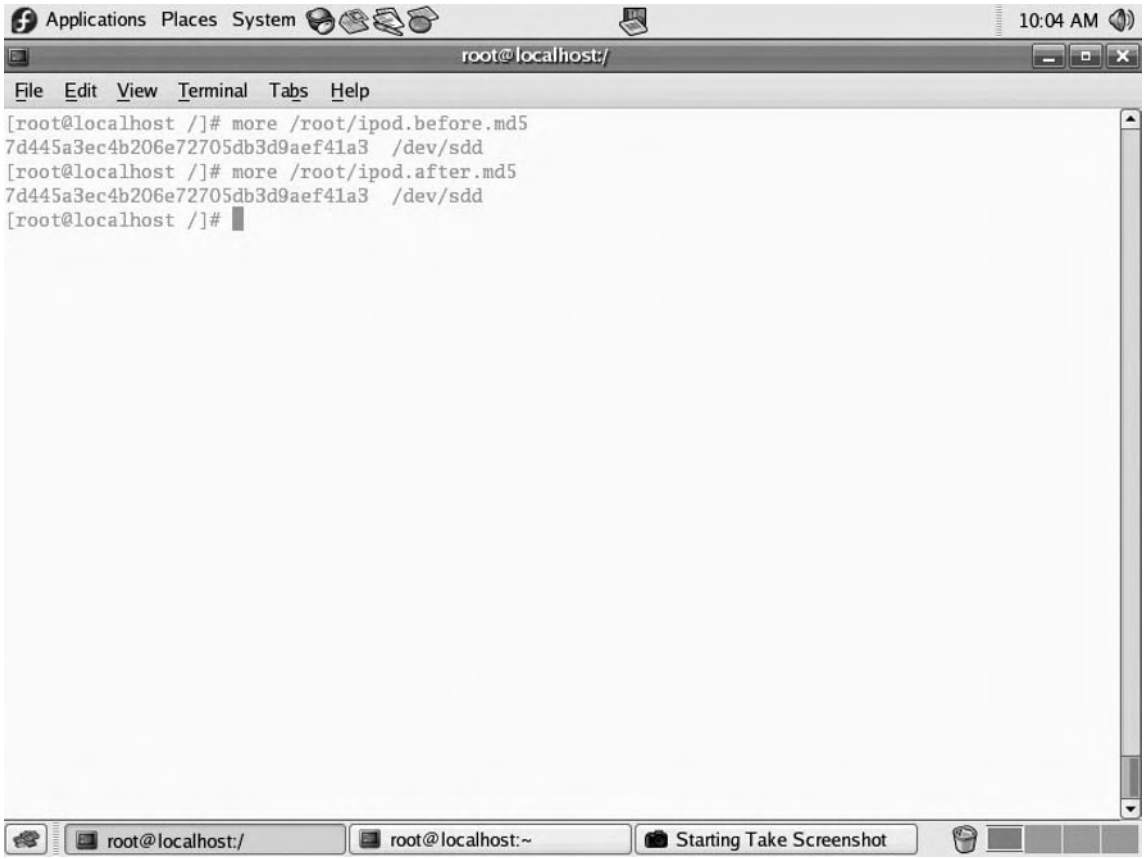
After the image is complete, perform another hash to ensure that the data has not been changed (see Figure 20.6).

Figure 20.6 Performing Another Hash



The next step is to compare the two hashes (see Figure 20.7).

Figure 20.7 Comparing the Two Hashes



As the previoius example shows, the before and after hashes of the iPod device are the same, which means nothing on the evidence drive was altered. Additionally, a hash of the forensic copy should be made to ensure that the hash of the image file is the same as the has of the iPod. This proves that the image contains the exactly same data as the iPod and the "dd" of the drive worked correctly.

Registry Keys

If you are using a Windows-based imaging tool such as Guidance Software's EnCase, you can use a key in the Windows Registry to write-block a USB device that is plugged into a forensic tower. This will keep Windows from writing to evidence. Doing a Web search on "write blocking USB device" will give further information on the steps necessary to carry out this procedure.

It is important to remember that using Linux or a Registry key edit for imaging is a last resort. It is always better to use a hardware write-blocking device. You can find many guides online that will detail the steps you need to follow if you choose to take this route. You can also go to www.windowsitpro.com/windowsstorage/Article/ArticleID/44380/44380.html.

Types of iPods

iPods come in many different physical and firmware versions.

The first generation of iPods became available to consumers in October 2001. They had a storage capacity of up to 10 GB. There have been many subsequent generations of the iPod. With each new generation, features became enhanced, including the addition of color screens, and video storage and playback capability. The storage capacity increased as well. Newer iPods can have storage capacities of up to 80 GB, using a Toshiba 1.8 hard drive.

The iPod Mini debuted in January 2004. The Mini was the first iPod available in various colors and was substantially smaller than other models. Storage capacity for the Mini was up to 6 GB, using a 1-inch Hitachi Microdrive.

The iPod Nano was the new version of the Mini. It was even sleeker and smaller and came in either black or white. Current Nano models have a storage capacity of up to 8 GB, using flash memory. The Nano has the ability to store and show digital photographs and video via a color screen.

The iPod Shuffle appeared in January 2005. The Shuffle used flash memory instead of a hard drive. The first Shuffle was smaller than a pack of gum. Unlike the other iPod models, it had no LCD display. The second-generation Shuffle was even smaller than its predecessor.

File Types Supported

Currently, iPods support the following file types: Advanced Audio Coding (AAC), Protected AAC, MPEG Audio Layer III (MP#), Variable bit rate MP# (MP# VBR), Audible Audiobook, Apple Lossless, Audio Interchange File Format (AIFF), Windows Audio, and Compact Disc Digital Audio, JPG, JPEG, TIF, TIFF, GIF, PNG, BMP, PSD, SGI, MPEG-4, and H.264.

File Systems

A file system is what organizes a computer's operating system. We discussed file systems at length in previous chapters.

The iPod uses two standard file systems: Microsoft's FAT32 and Apple's HFS+. The FAT32 file system is compatible with Apple Macintoshes and Windows PCs. HFS+ is writable only with Macintoshes. If a user has an iPod formatted with FAT32 and both a Macintosh and a Windows-based PC, he can read and write to the iPod using both file systems. Such a user can also write and read to the iPod using Linux.

The iPod is essentially a storage device and you can configure it to use almost any file system. I have used the extended 2 and 3, as well as FAT16 file systems on my iPods.

"Hacking Tools" and Encrypted Home Directories

At conferences, I like to show agents some worst-case scenarios. One of the things I like to show them is an iPod Shuffle that I have manipulated to act as a devious device capable of malicious activities. I repartition the hard drive so that there is enough capacity to install a bootable Linux distribution that contains various "hacking tools," including the popular Metasploit. The iPod control folder is left intact, along with all of the other folders needed for the iPod to function normally. I start the session by showing them the Shuffle connected to its iTunes library and playing music normally through speakers. I then plug the iPod into another machine and demonstrate how to boot it into Linux without touching the host machine. I show them how I can use Metasploit or another hacking tool to break into another machine and access data on that machine. I then explain to them how the host machine is never touched while I am hacking into it using the iPod and that all evidence of these activities is going to be found only on the iPod. The point of this exercise is to show that even if an iPod appears to act like it should, it may not in fact be what it seems.

Mojopac is another hacking tool for use with the Windows operating system. Mojopac allows a hacker to use an iPod as a virtual Windows desktop. Plugging an iPod into the USB port on a Windows computer copies the applications on that computer's desktop and allows the iPod to become a working virtual machine. For more information or to purchase this tool, go to www.mojopac.com.

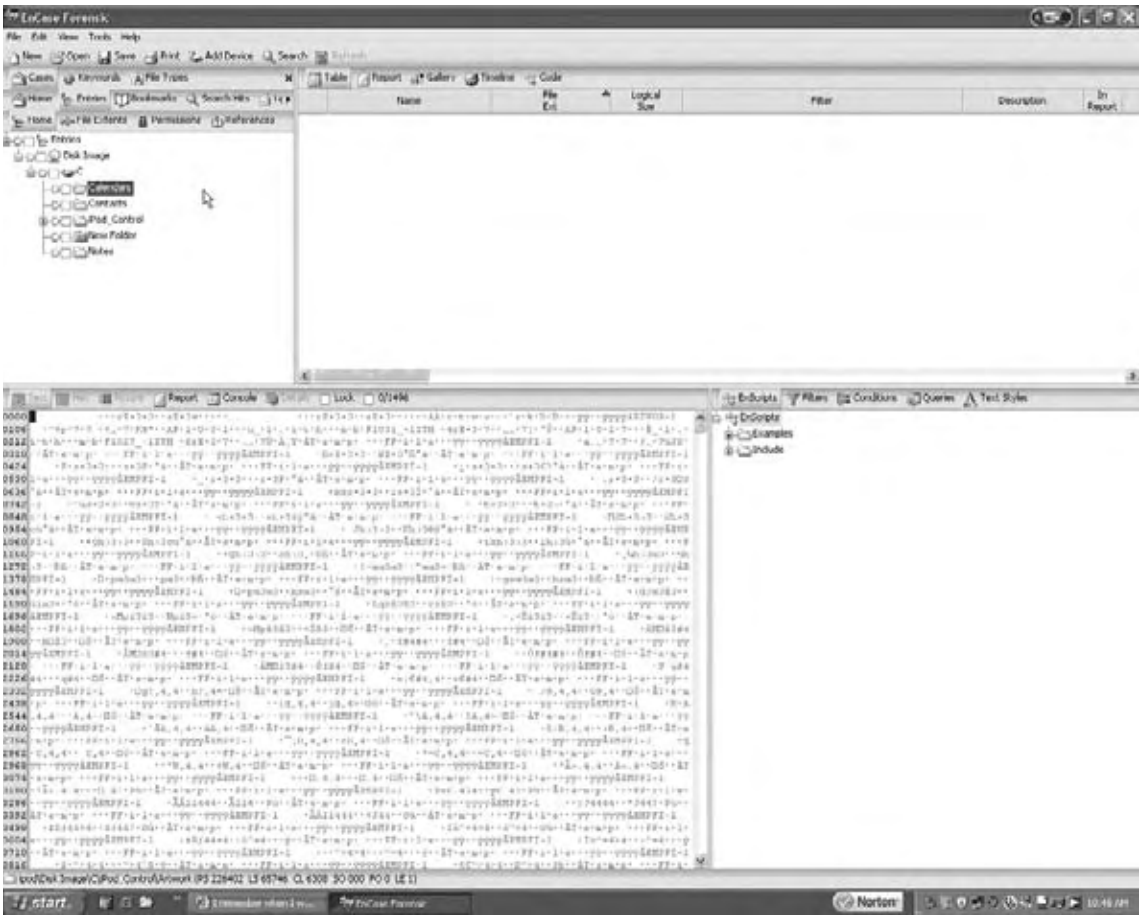
Another hacking technique using iPods is called "slurping." Slurping uses a tool called Slurp that captures documents, spreadsheets, and other files from the desktop of a computer using an iPod via the computer's USB port. This can be useful or malicious. For example, a malicious user could ask you whether she can use your computer to charge up her iPod using a USB port on your computer. Once the iPod is synced to the computer, Slurp captures all the documents and spreadsheets on your computer's desktop. The original article and code can be found at www.sharp-ideas.net/pod_slurping.php.

Evidence: Normal vs. Not Normal

When conducting an exam, forensic examiners need to know the distinction between normal data files and evidence that is not normal. Depending on the firmware and version of a particular iPod, there may be some variance in this determination throughout an analysis. For example, on older iPods, the song-naming convention displays the entire name of a song plus the music file extension, whereas on newer iPods songs are displayed with a four-letter code in addition to the file extension.

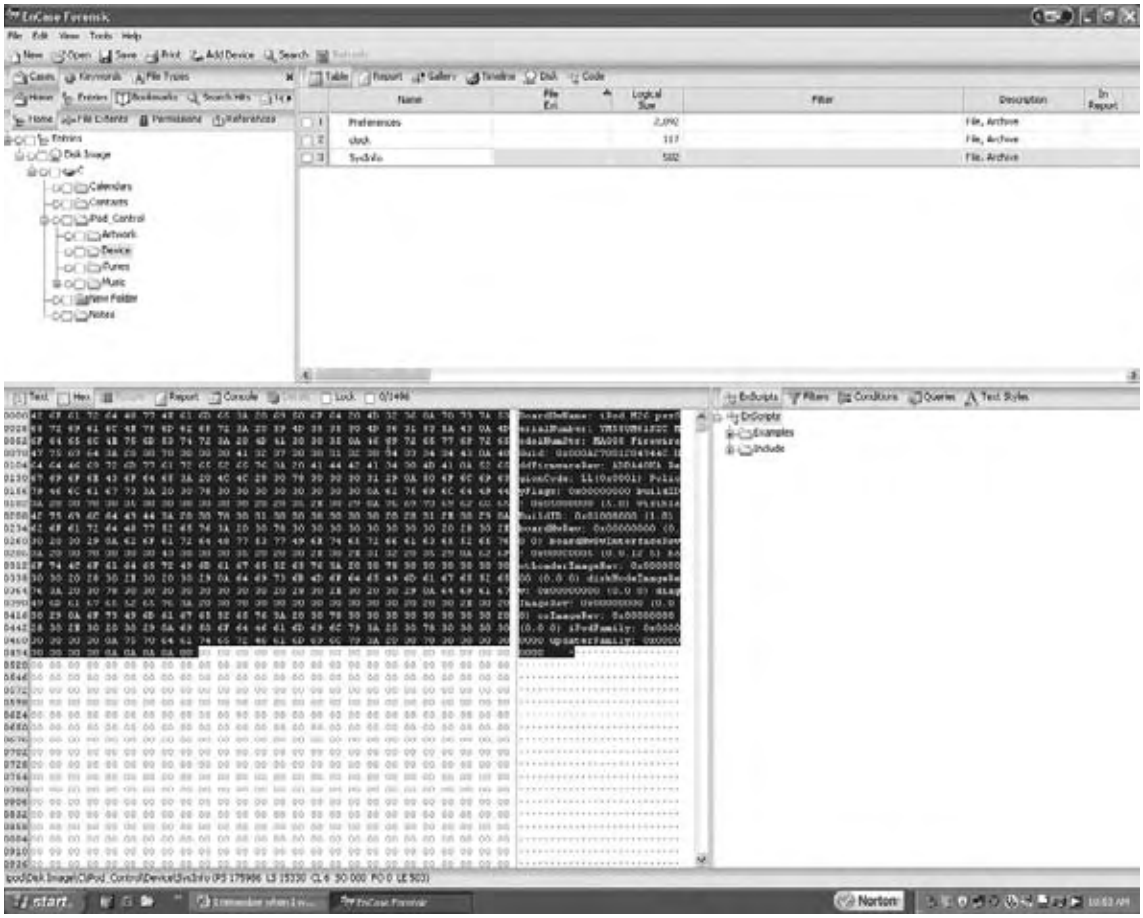
In Figure 20.8, you see the main directory structure of an iPod Nano, which contains the iPod_Control, Device, iTunes, Music, and Artwork main directories. The New Folder icon is not typical.

Figure 20.8 An iPod Nano's Directory Structure



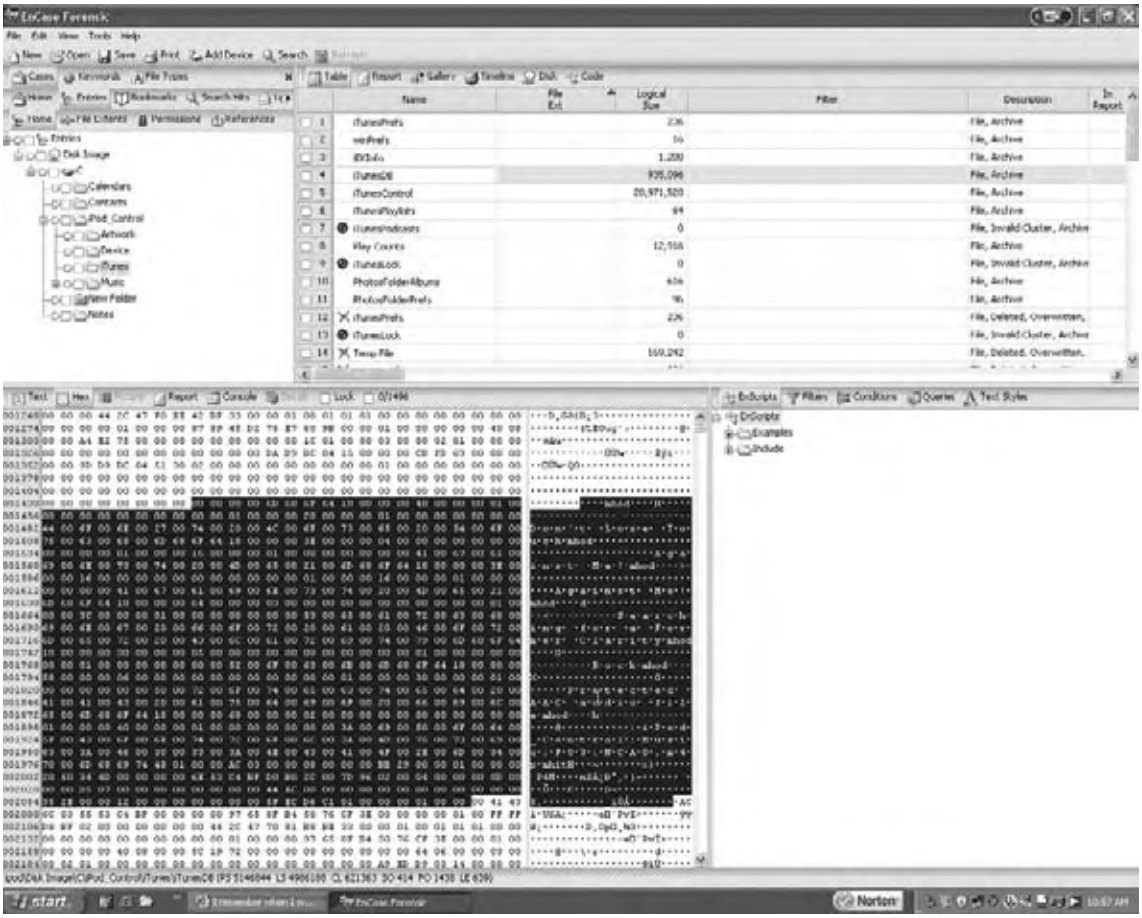
The Device folder contains files with some important information about the iPod, such as the firmware version and serial number such as you see in Figure 20.9.

Figure 20.9 This is the sysinfo output. You can see the serial number of the device.



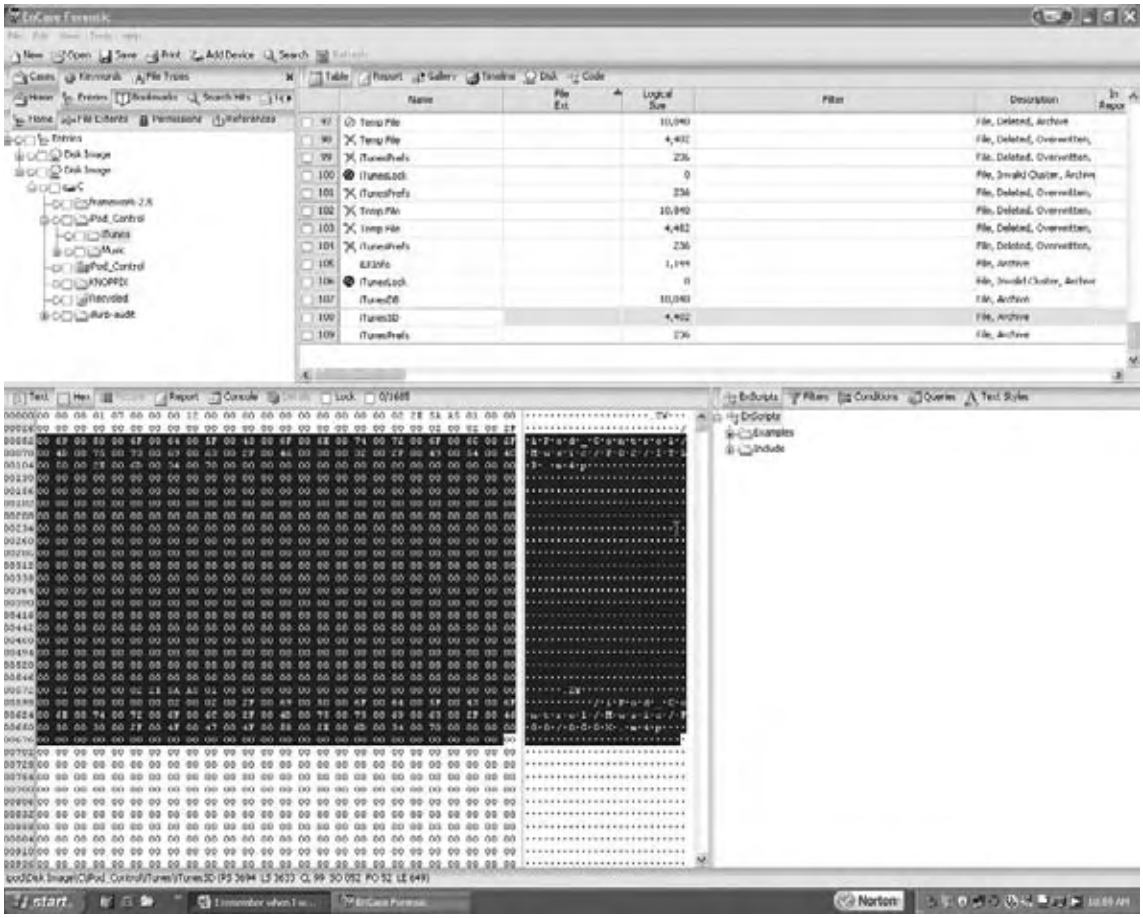
One of the files that forensic examiners note is the iTunes DB file, which provides information about music files, including their file type, music category, and the location on the device. This file is controlled by the iTunes software (see Figure 20.10). If a user manually moved a file onto an iPod, it would not be listed in the iTunes DB file. The file is found in the iPod_Control/iTunes directory.

Figure 20.10 An Example of the iTunes DB



The iPod Shuffle has a file called iTunes SD, which provides MP3 location and song title information. The shuffle is the only iPod which contains this file. There is an example in Figure 20.11.

Figure 20.11 An Example of the iTunesSD File



The iPod_Control directory is the control center of an iPod. It contains the Music and iTunes directories as you see in Figure 20.12. This is where all music files are stored by default. All music files are dispersed into various directories, each named F## such as you see in Figure 20.13.

Further investigation of the directories reveals the actual music files themselves. Newer versions of iTunes will condense MP3 or other digital music formated songs into four-letter codes followed by an extension, as shown in Figure 20.14, and as discussed earlier.

Figure 20.12 Inside the iPod_Control Directory

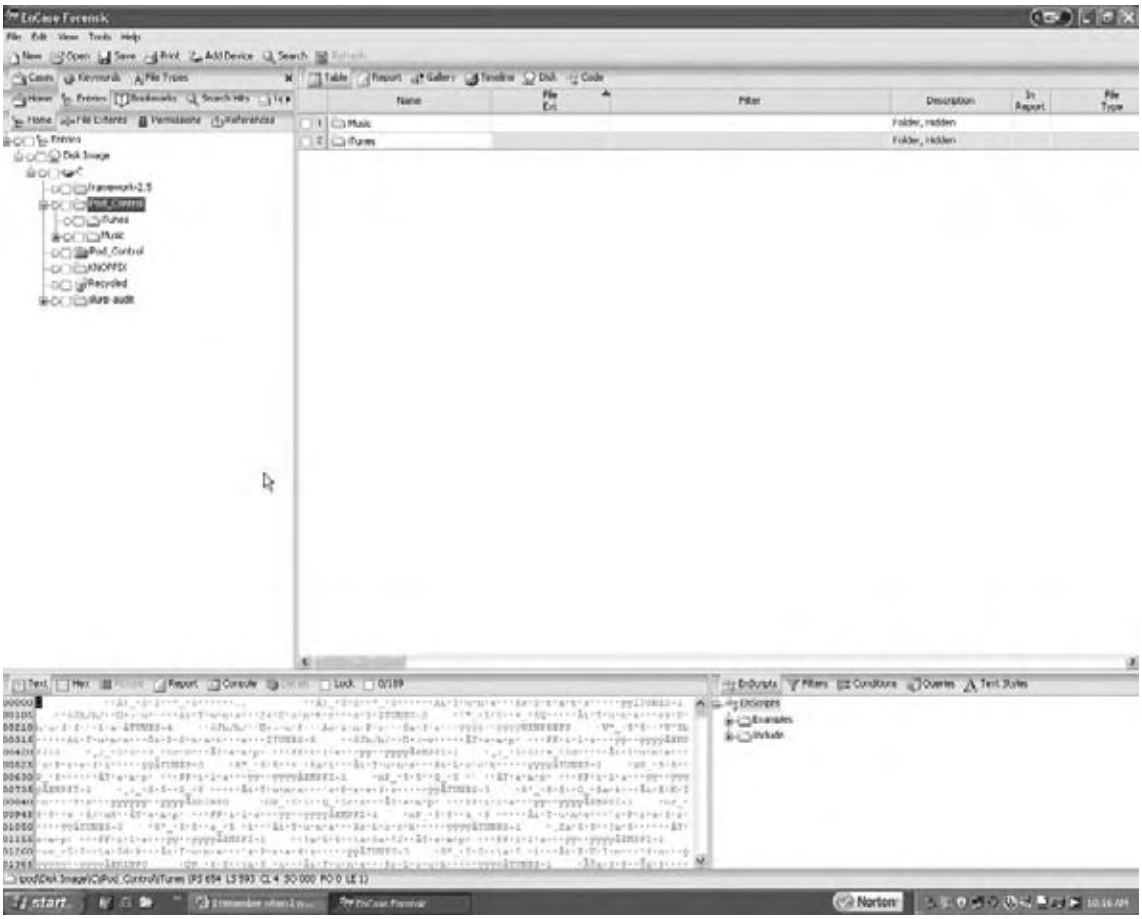


Figure 20.13 An Example of f## Music Directories

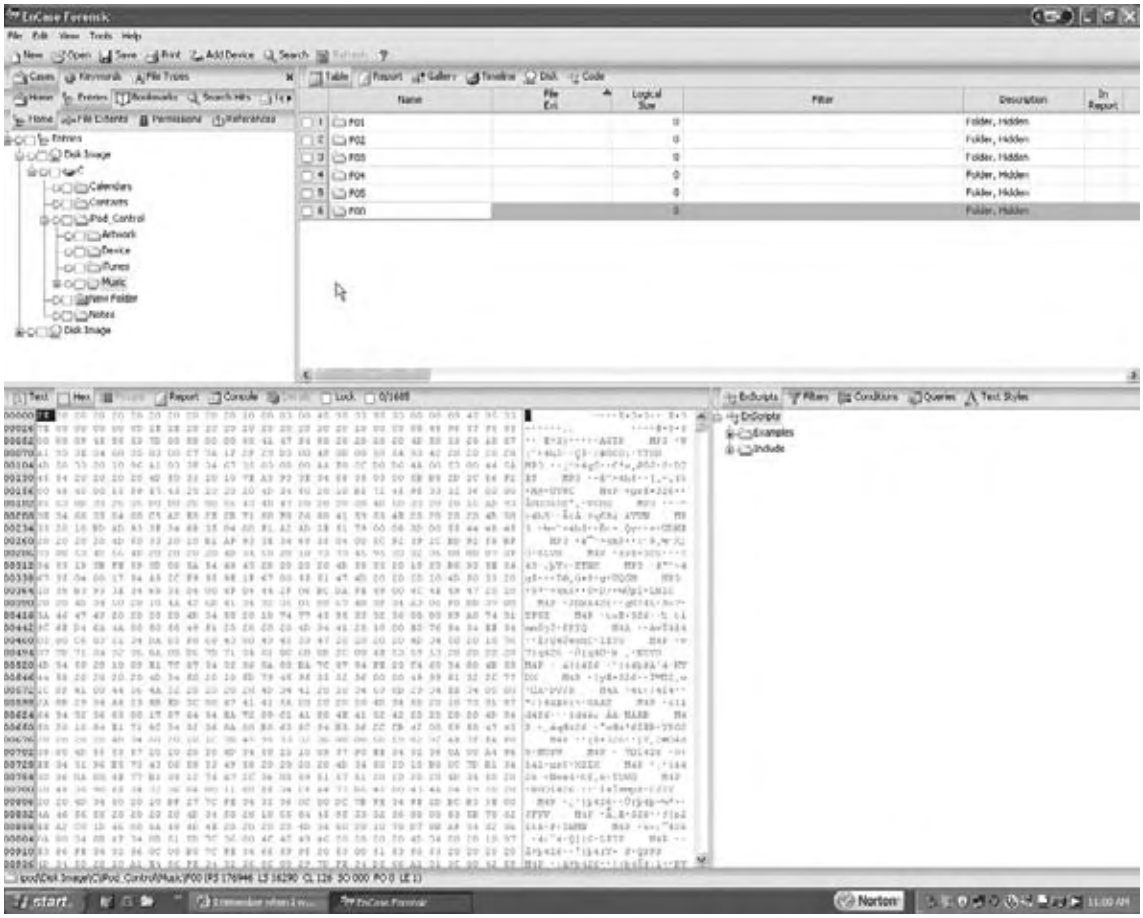
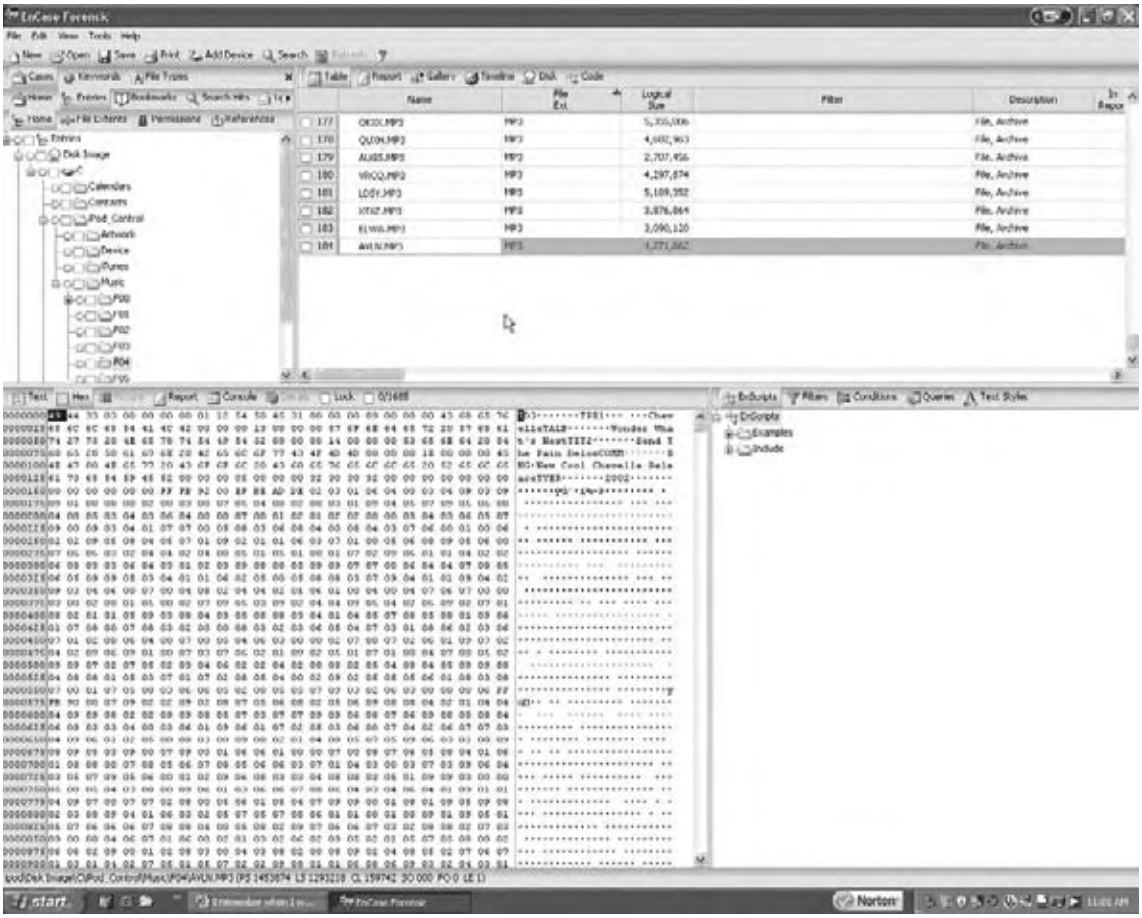


Figure 20.14 Example of an mp3 file on an iPod



When you are looking at digital photos or video files on an iPod, it is important to understand that the photos or videos themselves may be important evidence. The evidence could reside in plain sight on the iPod, or it could be hidden inside folders. For example, in a child pornography case, photo or video evidence might be in the default photo and video directories. Further investigation might be necessary to uncover hidden evidence.

Other directories are the Contacts, Podcasts, and Notes directories. Different versions of iPods have slightly different directories. For example, the iPod Shuffle has the Shuffle DB but does not have a picture-viewing directory. If there are photos in a Shuffle directory, those photos were placed on the device manually, not using the iTunes software.

Uncovering What Should Not Be There

Just because an iPod has been manipulated or changed from its factory configuration does not necessarily mean that there is suspicious activity going on. Many people like to change or hack their iPods. Sometimes it can be innocent, but other times it can be a telltale sign of malicious activity.

Suspicious items to look for are things such as mismatched file extensions. An example is a .jpeg file with a .mp3 extension. Most forensic tools are able to detect such discrepancies by using **signature analysis tools**. These tools find files that have a header that is different from the extension. You can configure most forensic tools to add custom file signatures.

Other suspicious items are hidden or improperly named files, which include files named something innocuous—for example, a photo that is named to look like an MP3 file. Additionally, files that should arouse suspicion could include those with blatantly outrageous names, such as "hax0r."

Too many partitions indicate that an iPod is not set to the factory default and should be looked at carefully. A file system other than the standard FAT32 or HFS+ installed on an iPod could indicate suspicious activity.

For example, the image in Figure 20.15 appears to contain a normal iPod directory structure. However, there are a few unusual items that bear notice, such as a Knoppix directory and the syslinux.cfg and ldlinux.sys files, which indicate that this iPod has some form of Linux on it. Also of note is the framework-2.5 directory. This directory contains the Metasploit hacking tool which can be found at www.metasploit.com. In this case, it happens to be Damn Small Linux (DSL), a very small, bootable version of the Linux operating system. There is also a slurp-audit directory, which is very suspicious.

Figure 20.15 A suspicious iPod

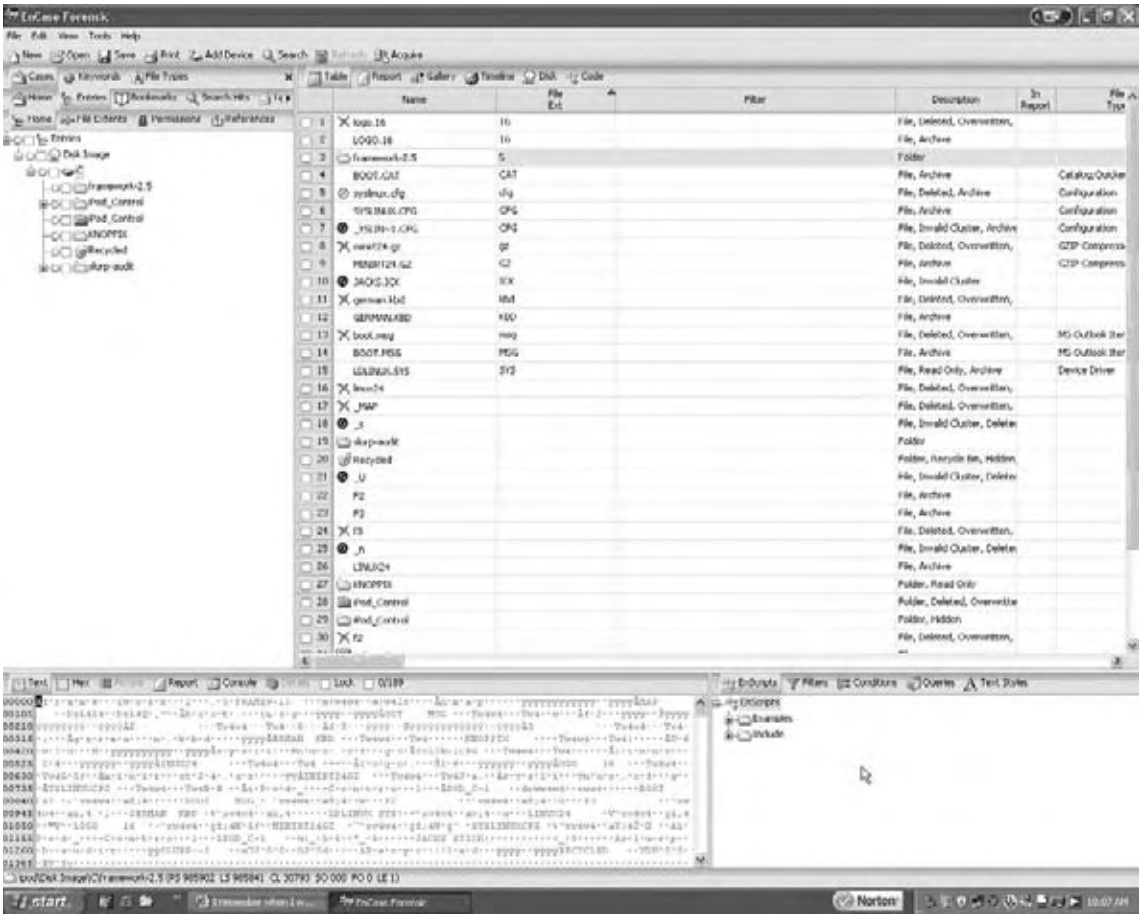


Figure 20.16 shows an example of an image disguised to look like an MP3 file. This is a simple and common way of attempting to hide evidence. Photos can be hidden in iTunes, which will load the disguised photo into its library and sometimes can be hidden as song files and placed into the F## directories to avoid detection.

The Music directory shows the standard F## directories. In Figure 20.16, opening one of these directories reveals two deleted songs that have the older iTunes song-naming convention of writing the whole song. This shows that the user had a previous version of iTunes. Additionally, the file named Hidden.mp3 is suspicious because it is not using the correct naming convention of either the old version of iTunes or the new one. Also it is named Hidden.mp3 which is a suspicious name in of itself. It is also much smaller in size than the

other regular music files, which indicates that something is wrong such as you see in Figure 20.16. In Figure 20.16, there are two other files that are smaller in size than regular music files typically are. These files might be image files that a crafty user has attempted to hide as music files.

Figure 20.16 Possible Suspect Music Files

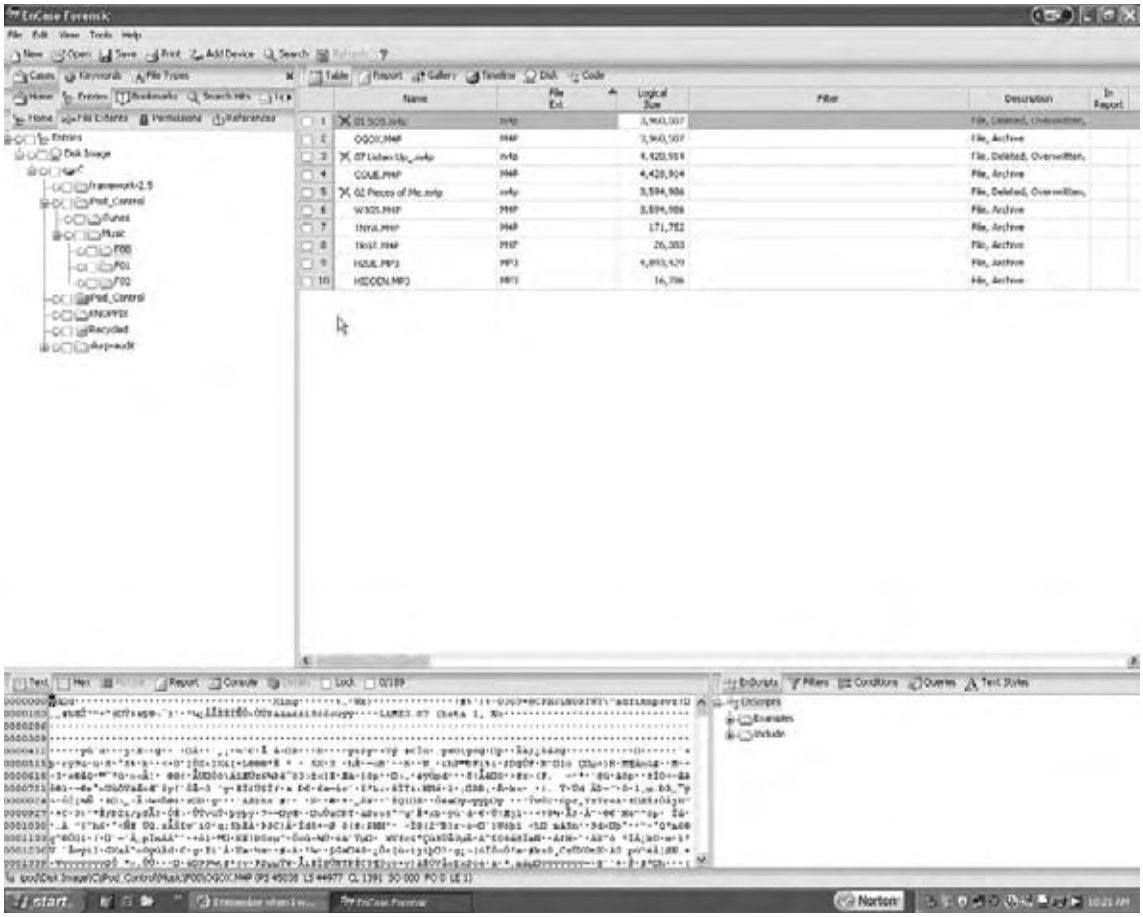
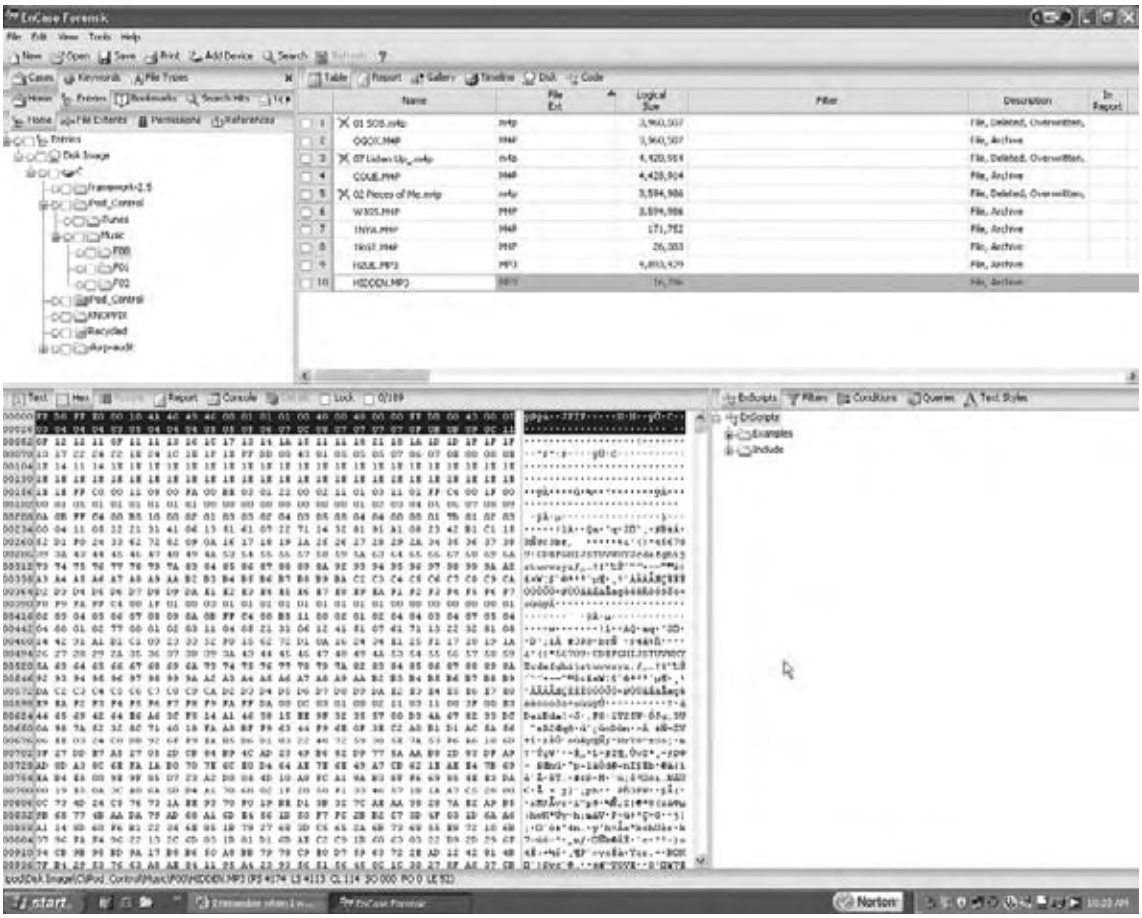


Figure 20.16 is an example of the inside of an F## directory.

Looking at the hidden.mp3 file using a hex editor shows that the JFIF file header indicates that this is not an MP3 as it should be (see Figure 20.17).

Figure 20.17 Music file with suspect header



Another tactic that users employ to disguise files is to insert text within a music file. A hidden message such as "The cow jumps over the moon at noon" might be inserted into an MP3 file. The MP3 will still play normally, which makes it difficult to detect. In this case, the best way to detect text within an MP3 file is through keyword searches. It is also possible to get hashes of songs from Apple and compare them to the song hashes on the suspect device. This will not show up on signature analysis because the actual file header will still match its extension.

Yet another way to hide photos is to make them cover art. With the color iPods, users can match cover art to music files. Default cover art is often included in songs purchased from iTunes. There are also Web sites that have current cover art. Users have the option of changing the cover art to suit their preferences, making it a good place to hide bad photos.

Analysis Tools

Forensic examiners can use almost any forensic tool that supports FAT32 or HFS+ for analyzing iPods, including Guidance Software's EnCase, AccessData's FTK, Brian Carrier's Sleuth Kit, and Paraben's P2. All of these tools utilize similar functionality to carry out analysis. All have a relatively intuitive user interface. The Sleuth Kit is primarily for more advanced users and runs only on the Unix/Linux platform. All of these tools are commercially available; the Sleuth Kit is downloadable free of charge. Not all of these tools will support the HFS+ file system, so users may be limited by particular file system parameters. All of these tools are capable of rendering image files and text files, and they have keyword search capability. A forensic examiner would use these tools in the same manner as he would in performing a static hard-drive analysis.

Summary

This chapter of the book introduced and explained the file structure of the iPod and showed how evidence can be hidden within the iPod.

In the field, I have seen everything from the standard iPod that stores only songs to completely tricked-out iPods running dual-booted operating systems full of exploits and hacks.

From a law enforcement standpoint, it is very important that search warrants specify that all data storage devices, including iPods, should be acquired as potential evidence. The iPod might be the sole source of evidence that makes or breaks a case.

Index

A

Access Data's FTK, 268
accreditation
  digital forensic laboratory, 133–134
  enterprise digital investigative/electronic discovery capability, 105
Active@ Data Recovery Software, 351
Active@ Partition Recovery, 370–371
AIM Password Decoder (password recovery tool), 189
Aircrack-ng, 583–586
alerts overview report, 539–541
alerts report, 536
Allin1 tool, 433–436
Alternate Data Streams (ADS), 505
American Society of Crime Laboratory Directors/LAB
  data forensics facility certification, 17
  model for facility operations, 18
anti-forensics, 73
AOL e-mail program, 268
application stupidity, 314–315, 330–331
ASCLD/ LAB. *See* American Society of Crime Laboratory Directors/LAB
Auditor Security Collection, 394
authentication
  detecting problems, 541
  as evidentiary value, 300–301
  password security, 637
  tracking failures, 542
  and watermarking, 632–633
Autopsy Browser, 515–516, 518–520
Autoruns (forensic software), 199
awareness-level training, 285–286

B

BackTrack
  acquiring images, 398–399
  case studies, 448–451
  forensic analysis tools, 413
  forensic tools, 395–397
  imaging information, 404–405
  and Magicrescue tool, 445
  and pasco utility, 424
  penetration-testing tasks, 581
  Rootkit Hunter tool, 426–427
BackTrack Mounting Drive, 397
BackTrack Security Collection, 395–396
backup and archiving managers, 161
Bart's PE Builder, 506
Basic Input-Output System (BIOS), 463–464
BestCrypt Encrypted Volume, 249
BestCrypt process, 251, 254
best evidence rule, 301
BlackBag Forensic Suite. *See* Macintosh Forensic Suite
Blackberry, 653–657
Black's definition of evidence, 269–270
blogging, 327
Booby traps, 376
boot disk, 256
boot up process. *See* system startup
brute force
  activity, 547–548
  attack, 544–546
BSSID. *See* MAC addresses
bulk demagnetizer, 340

C

cache monitor, 504
Captain Nemo, 523–524
Carbon Copy Cloner (CCC), 476–478
case reporting, 563–564
certification of personnel, 309
certification programs, 19
CFC (IEC-Council Certified Hacker Forensic Investigator), 105
29 CFR 1910.120(q)(6), 283–284
chain of custody, 562
chain of custody (evidence preservation), 7–8

chat programs, 293
chat sessions, 313, 327
chemical suppression methods, 24
command-line file-carving tool, 443–444
command line interpreters, 366–369
Command Prompt, 335–336
Command Prompt Delete, 336
comparing-Two Hashes, 671
Completed dd Function, 669
compressed files, 361–362
compression, 246
compromise, 564
computer crimes, 312–314
computer-facilitated crimes
  forensic investigation stages in tracking, 6–7
  law, 13
  reasons for, 4–5
computer forensics
  certification programs, 19
  computer evidence, 2
  digital evidence examination, 3
  flaws and risks
    modes of attack, 2
  history of, 3
  misconception, 244
  rules of evidence, 7
computer forensics laboratory, 13
  abatement strategies, 27
  accreditation of, 17
  acoustic balancing, 28
  anti-static flooring and sprays, 28
  auditing, 18
  business management and financial profitability, 14–15
  common office technologies, 31
  data destruction practices and tools, 40–41
  data handling strategies
    CD/DVD hardware solutions, 42–43
    for data storage and preservation, 41
    hard disk and magnetic tape, 42
  data storage space requirements, 39
  data storage technologies, 15
  design plan, 15–16
  desktop investigative gear, 38–39
  electrical demand of, 25
  EMI shielding, 28
  evidence locker security, 29–30
  evidence storage space, 20–21
  facilities build-out elements, 19
  fire protection plan, 21
    chemical suppression methods, 24
    class A/C hazard, 22
    gas agent suppression systems, 23–24
  forensics examination space, 20
  general ambience of, 30
  hardware duplication platforms, 35–36
  humidity management system installation, 28
  HVAC design, 27
  LAN/WAN load, 25
  in law enforcement and government agency, 16–17
  local workspace load, 25–26
  modes of operation, 14
  network facilities, 21
  personal workspace design, 31
  security, 29
  space planning, 20
  spatial ergonomics, 30–31
  standard operating procedure, 17
  tools, 32–33
  work area considerations, 31–32
  write block field kits, 34–35
computer forensics. *See* digital forensics
Computer Incident Response Suite, 192
computer security, 2
Convenient message viewer, 610
Coroner's Toolkit (TCT), 188
corporate world information infrastructures, 63
cost of noncompliance, 117, 119
cover generation techniques, 621
covert channel, 626
cyber
  See also computer crimes
  attacks, 6
  crime, 3–4

D
Darik's Boot and Nuke (DBAN), 185
data
  acquisition tools, 374–376
    features, 377–380
    types of, 381–383
  carving, 441–442
  cataloging, 148, 168–169
  destroying, 339–340
  destruction practices, 40
  duplication, 385–386
  hiding, 622–623
  in Linux, 386–389
  permanently destroyed, 339–340
Data Elimination Suite, 192
data forensics facility. *See* computer forensics laboratory
DataLifter (forensic software), 193–194
data mapping, 168–169
data objects
  as digital evidence, 270–271, 273
  explained, 266–267
  imaging, 295–297
  location of, 291–292
data recovery
  CD/DVD, 359–360
  compressed files, 361–362
  deleted images, 362–363
  hardware tools, 383–385
  Linux, 347–348
  Microsoft Office, 360–361
  tools, 349, 443
    Device Seizure, 182–183
    Directory Snoop, 184
    Forensic Sorter, 183
dcfl dd, 408–409
dd images, 403, 406
dd Options, 387–388
DDoS attacks, 560
dd_rescue, 411–412
decrypted file contents, 248
Decryption Collection Enterprise (password recovery tools), 189
degausser, 340
DEL/ERASE Switches, 336–337
deleted files, recovering, 334–335, 340, 348–349
deleted partitions
  command line, 365–369
  recovering, 334–335, 364–365
  recovery tools, 369–374
  using Windows, 364–365
Deming model, continuous quality improvement, 122–133
Denial-of-Service (DoS) attacks, 268, 560–561
Department of Justice Office of Inspector General's recommendations, 129–130
Detailed Alert Messages, 536
device configuration overlay (DCO), 59
Digimarc's Digital Watermarking tool, 632
digital crime scene model, 273
digital environments, 63–64
digital evidence
  collection, 273, 297–299
  extration of, 3, 11
  hash value of, 57
  IDE or Serial ATA hard drive, 59
  indexing, 66–67
  network investigations, 73–77
  non-contamination of, 3, 11
  procedure for establishing integrity of, 57
digital forensic certifications, 105
digital forensic laboratory
  achieving clients and implied needs, 120–121
  objectives for accreditation of, 133–134
  ownership, 136
  plan execution for developing, 128
  planning, 123–128
  quality
    assurance manuals (QAMs), 125–127
    control implementation, 130–131
    system requirements for, 119
  standard operating procedures (SOPs) for, 125–127

digital forensics. *See also* computer forensics
applications of, 12–13
case assessment and evidence preservation, 8
case studies, 448–451
crime scene information analysis, 7, 11
data recovery, 8–10
explained, 396
foremost tool, 455
hashing the source disk, 454
index.dat files, 452–453
interaction with corporate world, 63
legal issues, 12–13
Magicrescue tool, 456–458
methodologies
digital evidence collection and indexing, 66–67
identification phase, 65
litigation support, 65
presentation, 67
objective analysis and reporting, 62–63
principles, 54, 78
risk assessment, 55
tool kit contents, 9–10
tools, 394–395
training for, 62, 330
and volatile data, 485
digital information analysis, 298
digital investigative model, 272
digital investigator
hardware and software tools for, 59
objective analysis and reporting, 62–63
digital media analysis, 67
evidence analysis, 71–72
collection, 69–71
identification, 68–69
standard report format, 73
digital media, identifying, 275
Direct Connect model, live investigations, 255
Directory Scan utility, 473
disk cleanup, 338
disk encryption, 249–250, 279–280
DiskExplorer, 382
disk imaging tools, 186–187
DISKPART Commands, 367–369
distortion techniques, 621
documentation
of actions, 278
in digital evidence seizure, 299
need for, 275–276
PDA forensics, 650
reasons for, 527
DoD (Department of Defense) 5220.22-M. Commercial practices
DoD grade destruction practices, adherence to, 40
and PDWipe, 184
sterilization techniques, 41
and Wipe MASSter, 236
DriveLock IDE (forensic hardware tools), 235
DriveLook, 381
DriveSpy, 182
Drive wiping, 184
dry pipe system, 22
dynamic addressing, 316

E

EC-Council Certified Hacker Forensic Investigator (CFC), 105
e-discovery engineering, 148–149
cost-cutting measures, 151–152
and Federal Rules of Civil Procedure (FRCP) rules, 150–151
electronically stored information (ESI)
forms of, 154
identifying sources of, 152–153
locations of, 153–154
electronic discovery
costs, 151–152
information lifecycle managers (ILM), 159
and interviewing phase, 166–167
legal and IT team considerations for, 156–161
process for litigation, 162–165
requests for, 162–165
storage and federal rules and requirements, 149–151
tools for, 172–173
e-mail
archive card example, 597
archive types, 594, 598–599

e-mail (*Continued*)
attachments, 600
communication, 326
deleted data in, 601
examination tools, 600–601
functions, 594
ingredients, 597–598
managers, 159–160
phishing scam, 312
recovery of deleted, 613
terminology, 592–594
e-mail examiner (forensic software), 196–197
Encase, forensic data and analysis program, 195
encrypted file systems, contents of, 247
encrypted hard drive, forensic image of, 250
encryption, 279–280, 573–574
enterprise content management (ECM), 158
enterprise digital investigative/electronic discovery capability
accreditation, 105
administrative considerations for, 90–91
business case for, 86–87
certification, 104–105
costs for, 87–88
digital investigative team, 84–85
electrical and air conditioning requirements, 110–111
executive support for, 89–90
facility, 107–108
funding, 101–102
identifying requirements for, 85–90
identifying team resources, 105
digital forensic software, 106
hardware and storage, 106–108
write blocker, 108
need for developing, 84
organizational requirement for, 85
law enforcement agency, 86
legal department, 85–86
physical and logical security access controls to, 110
policies and standard operating procedures, 91–101
resources for, 89
time considerations, 88–89
training, 103
Ettercap, 587
evidence
admissibility, 301
authentication, 300–301
checking Internet Explorer, 503–504
corruption, 527
definition, 268–270
integrity, 562
and iPods, 673–674
locker security, 29–30
location, 484–486
logs, 56
searching for, 555
seizure, 267–268
volatility, 562–563
evidence seizure
best practices examples, 273
in child pornography, 296
digital, 267–268, 284
example, 289–291
labeling, 276
limiting factors, 278–282
minimizing, 300
nondigital, 303
other options, 286
and response to victims, 288–291
from running computer, 292–294
specialists, 283
Evidor, 181
Extensible Firmware Interface (EFI), 470

F

Failed Logon EventIDs, 543–544
failed logons
identifying, 543–544
listing, 542
Faraday enclosure, 43–44
Farmer's Boot CD (FBCD), 524–525
FastBloc LE, data acquisition and write protection device, 195
FavURLView (forensic software), 189–190
FBI laboratory and quality system, 129–130

FDISK command, 482
fdisk -l Command, 401–402
Fdisk output, 665
FDISK Switches, 366–367
Federal Rules of Civil Procedure (FRCP)
  changes in, 150–152
  and digital information management, 84
Federal Rules of Criminal Procedure (FRCP)
    Rule 41, 270
Federal Rules of Evidence Rule 901
    (28 U.S.C.), 300
Fernico FAR system, 43
file(s)
  access attempts auditing, 549–550
  carving, 441–442
  cleaning operation, 251
  deleting, 335–339
  deleting permanently, 344–345
  headers and extensions, 154
  hiding, 555–556
    See also steganography
  integrity checker, 185–186
  moving, 337, 339
  signature, 442
    header, 443
    search, 96, 101
  slack, 495, 527
  systems
    auditing, 549–550
    with Captain Nemo, 524–525
    examination, 496
FileSpy, 473–474
FireWire DriveDock and Lockdown (forensic
    hardware tools), 234–235
first responder, 284–285
first responder training, 304–305
f## Music Directories example, 679
foremost tool, 443–445, 455
forensic
  analysis, 413
  computing, 3
  investigator, 6–7
forensic hardware, 239–240, 392
  hard disk write protection tools
    DriveLock IDE, 235
    FireWire DriveDock and Lockdown, 234–235
    Image MASSter 3004 SATA, 239
    ImageMASSter 3002SCSI, 238
    Image MASSter Solo-3 IT, 236–237
    NoWrite, 233–234
    Wipe MASSter, 236
    Write Protect Card Reader, 235
forensic software, 46, 201–233
  ad hoc scripts and programs, 50
  Computer Incident Response Suite, 192
  Coroner's Toolkit (TCT) and Tctutils, 188
  Data Elimination Suite, 192
  DataLifter, 193–194
  e-mail recovery tools
    Autoruns, 199
    HashDig and PowerGREP, 200
    Network E-mail Examiner and Oxygen
      Phone Manager, 197
    Reverse Engineering Compiler (REC),
      200–201
    SIM Card Seizure, 198–199
  Encase, forensic data and analysis program, 195
  enterprise forensic tools, 49–50
  Evidor, 181
  File Integrity Checker
    File Date Time Extractor (FDTE) and
      Forensic Date/Time Decoder, 186
    FileMon, 185–186
  file systems, 47–48
  and hardware, 43
  Image Master Solo and Fastbloc, 194–195
  investigative platform, 48–49
  Ltools and Mtools, 187–188
  Maresware, 190
  NTI Secure ToolKit, 192
  operating systems, 47
  packages, 414
  password recovery tools
    AIM Password Decoder, 189
    Decryption Collection Enterprise, 189
    LOphtCrack application, 188–189

forensic software (*Continued*)
PHOTORECOVERY, 191
preview software, 292, 302
ProDiscover DFT, 191
R-Studio, 193
SnapBack DatArrest (disk imaging tool), 187
special focus tools, 49
Stealth Suite, 192
TextSearch Suite, 192
Visual TimeAnalyzer, 178–179
WinHex Specialist Edition, 191
X-Ways Forensics, 179–181
Forensic Suite Toolbar, 472–473
Forensic Toolkit (FTK), 194
forensic tools
for analysis, 413
in Auditor Security Collection, 394
in BackTrack Security Collection, 395
for digital evidence collection, 297–298
opensource, 448
pasco utility, 424
Forensix Project, 522
forms chain of custody, 56
fport, 494
freeware tools, use of, 392
fsstat Command-Line Tool, 431
FTK (Forensic Toolkit), 194
FTK Imager, 378
fully processed MS Exchange file, 611
funding, digital investigative/electronic discovery capability, 101–102

G

gas agent suppression systems
inert gases and fluorine compound gases, 24
laboratory fire control, 23
gateway, 319
government information infrastructure
IT security staff, 63
official investigation, 64
GroupWise mail, 596
GRUB booting, 469
Guidance Software's EnCase Forensic examination suite, 40

H

Hacker Defender
and network analysis, 261
in psychical memory, 260
hacking
methodology, 554–555
process, 554–555
tools, 673
hard disks, 39
hardware
duplication platforms, 35–36
seizure, 267
and software selection process, 59–60
hash analysis search, 95, 100
hash code, 57
HashDig and PowerGREP (e-mail recovery tools), 200
hashing, 9, 405
HeaderBuilder, 474–476
HELIX Incident Response, 256
Helix Live Response, 486–487, 489–492
HMG Infosec Standard No 5, 41
home digital environment, 64
hostnames, 320
host-protected area (HPA), 59, 430–431
hotspots, 322–323
Houston Police Department Laboratory (HPD-LAB) issue, 118–119
Hydan tool, 623–624

I

Identity Theft, 312
IECookiesView, 503
IE History Viewer, 503–504
Image MASSter 3004 SATA (forensic hardware tools), 239
ImageMASSter 3002SCSI (forensic hardware tools), 238
Image MASSter Solo-3 IT (forensic hardware tools), 236–237
Image MASter Solo, system integration and MIS tool, 194–195
image steganography, 619

imaging and hashing, iPods, 663–664
imaging information, 294–295
  in Linux, 514
  using BackTrack, 404–405
  using dcfl dd, 410–411
  using NetCat, 407
IM managers, 159–160
incident response, 564, 567
individual account in MS Exchange file, 609
information hiding, 626–627
  See also steganography
  in Blackberry, 656
information infrastructure
  academic, 64
  in corporate world, 63
  government, 63–64
  internet and home, 64
information lifecycle managers (ILM), 159
information management system, 20–21
insider and external attacks, 6
instant messaging, 293, 313, 327
International Association of Computer Investigative Specialists (IACIS), 61
International High Technology Crime Investigation Association (HTCIA), 61
International Organization for Standardization (ISO), 122
International Society of Forensic Computer Examiners (ISFCE), 61
Internet Explorer, 64, 503
Internet Service Providers (ISP), 315–318
interpersonal communications, 325–327
Intrusion Detection/Intrusion Prevention System (IDS/IPS), 296, 533, 536, 580–582
intrusion process, 554–555
IP address, 315–320, 531–532, 536, 538–539
ipconfig/all Command, 317, 321
iPod_Control Directory, 678
iPod Nano's Directory Structure, 674
iPod Nano's sysinfo output, 675
iPods
  analysis tools for, 685
  child pornography case, 662
  file systems, 672–673
  history of, 662–663
  imaging and hashing, 663–664
  suspicious, 681–684
  types of, 672
iTunes DB example, 676
iTunesSD File example, 677

J

JPEGs recovery, 446–447

K

keyword search
  electronic discovery SOPs, 100
  for keylogger term, 259
Kismet, 582–583

L

laboratory analysis delays, 281, 283
Lattes, Leone, 3
law enforcement training, 282–284
LCP tool, 639, 641
legal documents, examples of, 270
LILO booting, 469
linguistic steganography, 622
Linux
  advantages, 508
  boot process, 513–514
  data recovery, 347–348
  dd program, 398–400
  directory path, 509–511
  disk forensics challenges, 514
  file system description, 508
  forensic tools, 515
  imaging information, 514
  and iPods, 665–670
  loading, 468
  Mount Command, 512–513
  primary directories, 509–511
  SMART, 520
live forensics
  boot disk, 256
  disk encryption, 249–250
  need for conducting

live forensics (*Continued*)
encryption usage, 247–250
evolution of enterprise network, 244
evolution of storage, 245–247
physical memory imaging
BestCrypt process, 254
memory, unencrypted document in, 253
ProDiscover IR Imaging Screen, 251–252
and postmortem forensics, 244, 256–257
live response, 76
Live System Information, 257
LiveWire, investigative software platform
Hacker Defender in psychical memory, 260
keyword search, 259
live system information, 257
running processes, 258
local storage archives, 596–597
logging requirements, 587
logical address. *See* IP address
Logon Events with No Logoff Event, 548
Log Parser
and failed logons, 542
file access attempts, 549
and logoff events, 547
and Snort, 534–535
uses, 551
LOphtCrack application (password recovery tool), 188–189
Lotus Notes, 595
Ltools application (UNIX tool), 187–188

M

MAC addresses, 320–321, 576, 584
MacDrive6/7, 478–480
MAC filtering, 576
Macintosh Forensic Suite, 472–475
Macintosh Options, 480
Magicrescue tool
case studies, 455–458
recipe files, 445–447
magnetic tape, 42
mail application, 605
mailbox archive, 598
malicious software, 260
Maresware, 522–523
Master Boot Record (MBR), 464–466, 482
Master Partition Record, 464–465
mboxgrep, 417–419
MD5 generator, 493
MD5 Hash, 295, 405–406, 475, 665
MD5 hashing algorithm, 57, 215
md5sum command, 405
media, size of, 279, 328
media sterilization systems
forensic hardware and software duplication tools, 39
memfetch, 419–423
Memfetch find, 424
Memory Dump Files using memtech, 421–422
memory dumps, 420–423, 497, 499
metadata fields, 155–156
Microsoft Office repair and recovery, 360–361
Microsoft Security IDs, 498
mobile device investigation, 648
modems, 316
"More" Command, 667
Mount Command, 403, 512–513
mp3 file on an iPod example, 680
MSDOS, loading of, 466–467
MS Exchange, 595
MTools application (UNIX tool), 188
Multi Platform File Manager, 523
Murphy's Law, 62
Music file with suspect header, 684
Myspace, 327

N

National Institute for Standards and Technology (NIST) program, 298–299
Nessus, 580
NetAnalysis, 190
NetCat, 388–390, 406
Netcat Options, 389
Network Address Translation (NAT), 532, 552
network
analysis, 261
archives, 613
explosion, 318–320

network (*Continued*)
facilities, 21
forensic and investigative softwares, 255
forensics, 554, 566
information-gathering tools, 532–533
intrusion, 268
Network E-mail Examiner and Oxygen Phone Manager (e-mail recovery tools), 197, 610
network interface card (NIC), 316–317, 320
network investigations, 73
entities, digital devices, and potential evidence, identification of, 74–75
evidence collection, 75–76
network trace evidence analysis, 77
Network Mapper l (Nmap), 579–580
networks, wireless, 570
access points, 573, 579–580
advantages, 572
basics of, 571–572
disadvantages, 572
penetration testing, 578
NIST (National Institute of Standards and Technology), 51
nondigital evidence, 303
NoWrite (forensic hardware tools), 233–234
NTFS file system, 505, 549
NTI Secure ToolKit (forensic software), 192

O

off-line tape discovery, 161
On Demand Connection model, live investigations, 256
on-scene responders
See also first responder
appropriate seizure method, 302
and evidence seizure, 272–273
imaging information, 294
level of knowledge, 299–300
options for, 267–268, 271
previewing information on-scene, 291–292
and prioritized ranking, 276
and Search and Seizure Manual, 286–288
Ontrack Data Recovery, 181–182
Ontrack PowerControls, 607–608
opening of data in MS Exchange file, 612
Open Shortest Path First (OSPF), 558
open source tools, 394, 396, 448
Open Systems Interconnect (OSI) model, 530
operating system artifacts, 72
Optical media and data preservation, 42–43
Osborn, Albert, 3
OSI Model layers, 530–531, 552
OS X kernel, 471
Outguess Tool, 629
Outlook Express, 601, 605
Outlook PST, 601–602, 613

P

Paraben's E-mail Examiner, 602–603
Paraben's PST Converter, 603
Partimage (Linux utility), 187
partitions, hidden, 626
pasco utility, 424–426
password
cracking, 637–641
security, 642–644
PDA devices mishandling, 652
PDA forensics, 648–650
PDA managers, 159–160
PDA Secure tool, 652
PDA Seizure, 652
PDCA cycle. *See* Plan-Do-Check-Act cycle
PDWipe, 184–185
penetration tests, wireless, 578
Penguin Sleuth, 520–522
Performing Another Hash, 670
personal computer, general seizure of, 272–273
Personal Digital Assistant (PDA), 648
PHOTORECOVERY (forensic software), 191
physical media, prioritizing, 275–276
physical memory imaging (live forensics)
BestCrypt process, 254
memory, unencrypted document in, 253
ProDiscover IR Imaging Screen, 251–252
physical seizure, 302–303

Plan-Do-Check-Act cycle, 122–133
P-5239-26 MFM/RLL standard, 41
policy
  chain of custody, 55
  for enterprise digital investigative/electronic discovery capability, 91–101
portable device forensics
  and data storage, 45
  Faraday enclosure, 43–44
  forensic toolkits, 46
portable enterprise systems, 38
portable forensics systems
  digital evidence analysis, 36
  laptop style, 37
  suitcase workstation style, 38
postmortem forensic analysis
  decrypted file, 248
  and encryption usage, 247
  files/processes, accessibility of, 260
  physical memory (RAM), 259
Postmortem *vs.* Live Forensics, 244
Power On Self Test (POST), 463–464, 470
preaction systems, 22
pre-built forensics systems, 39
Pre-Deployed Agent model, live investigations, 255
privacy concerns, 280–282
Privacy Protection Act (PPA), 280
processed files, 604
processing with forensic tool, 606
Process MS Exchange PRIV.EDB file, 608
ProDiscover DFT (forensic software), 191
ProDiscover IR Imaging Screen, 251–252
project scoping, 166
Pslist, 494
Psloggedon, 495
pull-the-plug, 277–278

Q

quality
  assurance manuals, 125–128
  audits, 129
  control, 129
  defined by American Society for Quality (ASQ), 117
  management principles, 122
  management system, 116
    and internal quality processes, 131
    international standards for establishing, 122
    ownership, 135
    quality audits, 129
    and quality manager (QM), 138

R

RAM dump, 292–294
RAM recovery, 292–293
RAM slack, 495
records and information managers (RIM), 158–159
recycle bin, 345
  damaged, 345–346
  and deleted files, 340–342
  properties, 342
  replacement tools, 358–359
  storage locations, 343–344
RegScanner, 497–498
Reverse Engineering Compiler (REC) (e-mail recovery tool), 200–201
RFC 3227, 485–486
RFID (Radio Frequency Identification) tags, 44
RIM device, 653
rm Switches, 347
Rock XP, 642
rogue access points, 580–581
Rootkit Hunter tool, 426–427, 429
routers, 318, 556–557, 566
  architecture, 557
  attack topology, 559
  hacking, 559
  hardening, 559
  investigating, 562
  wireless, 323
routing
  attacks, 561
  protocols, 558
  tables, 556–557, 561
Routing Information Protocol (RIP), 558
R-Studio (forensic software), 193

Rules on Criminal Procedure, 271
running services, 256
  hacker targets, 257
  and open network ports, 258

S
Scientific Work Group on Digital Evidence (SWGDE) and QAM documents, 139–140
Search and Seizure, 270
Search and Seizure of Computers and Obtaining Digital Evidence (Manual), 271, 286–287
SEARCH primer, 293
search warrant, 578–579
security policy violations tracking, 547
seizure
  process, 299–300
  protocols, 272
seizure methodology
  appropriate, 302–303
  current, 274
  digital evidence, 271–272
  example, 289–291
  fluidity of, 286
  steps, 275–277
  traditional, 273
  using minimization, 274
  warning, 277
server level archives processing, 606–607
server storage archives, 594
Service Set Identifier (SSID), 573, 576–577
SHA-1 hashing algorithm, 57, 95
shareware tools, 392
shred Switches, 347
shutdown method, 277–278
Sigverif.exe, 496
slack space, 399, 626
Sleuth Kit tools
SMART, 520
Smurf, 560
SnapBack DatArrest, 187
sniffers, 532–533, 586
snort, 534–535, 537
snow tool, 627–628
social networking, 327
software licensing, 50
software packages, forensic, 268, 414
SOP. *See* standard operating procedure
Spam Mimic, 620
spread spectrum techniques, 621
standard operating procedure, 17
  for creating bitstream image, 58
  for digital forensic laboratory, 125–127
standard operating procedures
  digital investigations, 92
    data analysis phase, 94–96
    data collection, 93–94
    data reporting phase, 96
    documentation, 93
  electronic discovery, 96–101
standards
  DoD 5220.22-M standard, 40–41
  HMG Infosec Standard No 5, 41
  P-5239-26 MFM/RLL standard, 41
static addressing, 316
statistical method techniques, 621
Stealth Suite, 192
steganalysis tools, 630–631
steganography
  attacks, 635–636
  categories of, 620–621
  classification of, 618–619
  detection, 633–635
  file systems, 625
  future of, 618
  history of, 616–618
  manipulation, 624–625
  techniques, 619–620
  tools, 627–631
  types of, 622–623
storage devices, 279–280
  and media, 328
  seizure methodology, 276, 286
  wireless, 325
StrongHold Bag, 44
substitution system techniques, 620

Suspect Music Files, 683
sustainment budget models, 102
Switch Description, 336–337
switches, 532
SYSINT process, 466–467
system scanner, 500–501
system startup, 462–465
system state backup, 504–505

T

Tableau T35i Combination Bridge, 34, 39
technical steganography, 622–623
TextSearch Suite, 192
The Sleuth Kit (TSK), 515–517, 685
  tools
    27 Command-Line Forensic Tools, 429
    and Allin1, 433–435
    command line, 432–433
    using, 428–431
Tool testing
  software and hardware, 50
  test script, 51
Traces Viewer, 503
training, digital investigations, 103
transform domain techniques, 620–621
tree view of MS Exchnage data, 611
trusted software, 76
Two Stage Loader, 469

U

Ultimate Boot CD-ROM, 506–507
UNDELETE Switches, 350
undelete tools, 350
  command-line tool, 351
  freeware, 353–354
  types of, 352, 355–358
user activity monitoring, 541
using Gifshuffle, 628

V

/var/log/rkhunter.log, 427
Vinetto, 437–441
virtual memory, 499–500
Visual TimeAnalyzer, 178–179
volatile data, 485–486, 562–563
volatile evidence, 10

W

wardriving, 323–324, 577
watermarking, 632–633
Web hosting/collocation providers, 76
wet pipe system, 22
White Canyon Software's WipeDrive, 40
WIGLE (Wireless Geographic Logging Engine), 577
Windows boot menu, 468
Windows CD-ROM bootable, 506–507
Windows forensic process
  advantages, 486
  evidence location, 484–485
  gathering volatile data, 485–487
Windows Forensic Toolchest (WFT), 493
Windows host, 484–485
Windows Registry, 497, 671
Windows XP, loading of, 467–468
Wipe MASSter (forensic hardware tools), 236
Wiping, 58
Wired Equivalent Privacy (WEP), 574
wireless attacks, 570–571
wireless networks, 322–324, 570–571
  access points, 573, 579–581
  advantages, 572
  attacks, 587–588
  basics, 571–572
  disadvantages, 572
  penetration tests, 578
  safety, 590
  security features, 573–575
Wireless Packet injection, 584–585
Word Extractor, 497
write blocker, enterprise digital investigative/electronic discovery capability, 108
write blockers, 33–34
Write Protect Card Reader (forensic hardware tools), 235

X

X-Ways Forensics, 179–181, 501–502